GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 10:44:58 ; Search time 11824 Seconds
(without alignments)
11385.611 Million cell updates/sec

Title: US-09-900-751-1
Perfect score: 3106
Sequence: 1 catggtagacggctgcccgg.....ttaaaaaaaaaaaaaaaaaa 3106

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

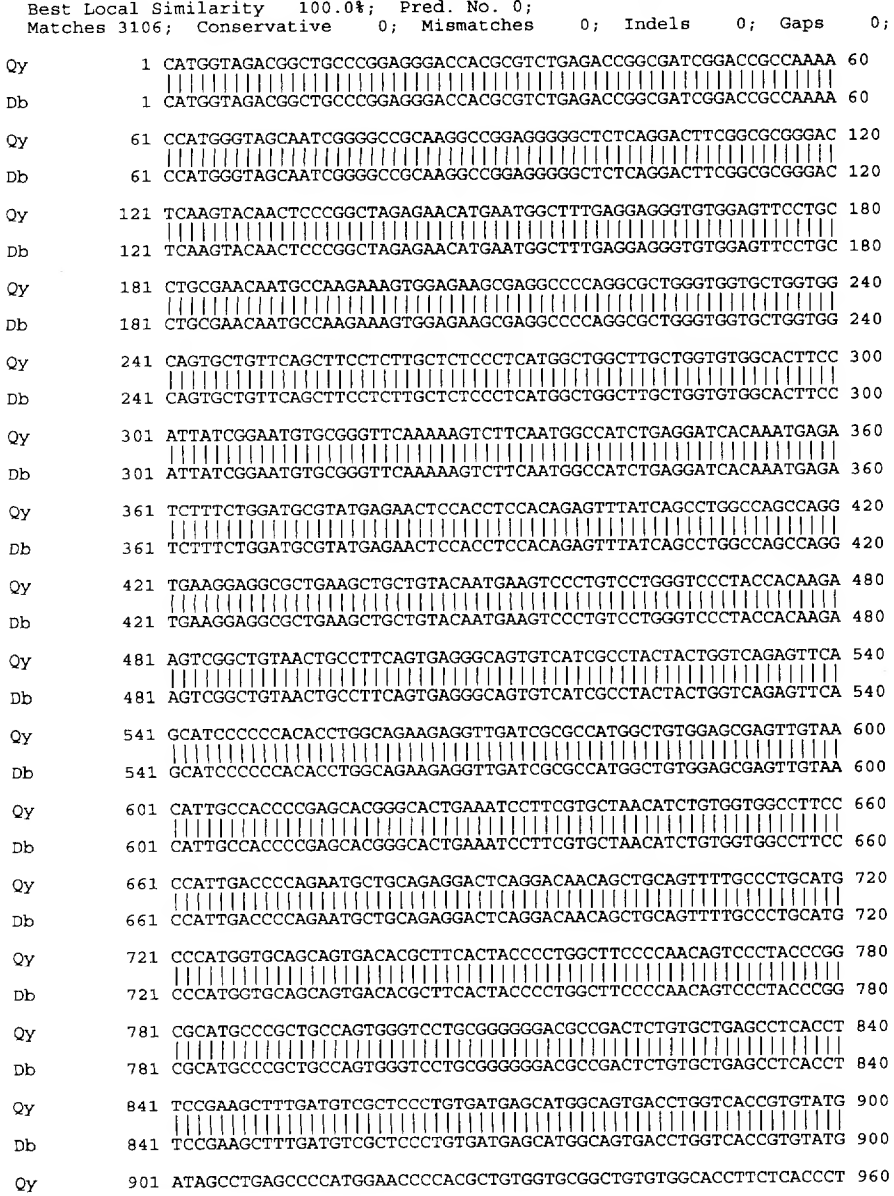
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3106	100.0	3106	6	AX395268	AX395268 Sequence
	2	3106	100.0	3106	10	AF042822	AF042822 Mus muscu
	3	3102.8	99.9	3248	10	BC005496	BC005496 Mus muscu
	4	2595	83.5	3174	10	AB049189	AB049189 Rattus no
	5	2313.6	74.5	2568	10	AB037898	AB037898 Rattus no
	6	1905.6	61.4	3273	9	BC030532	BC030532 Homo sapi
	7	1889.8	60.8	3149	6	BD274671	BD274671 Matriptas
	8	1889.8	60.8	3149	9	AF118224	AF118224 Homo sapi
	9	1884.8	60.7	3120	9	AF133086	AF133086 Homo sapi
	10	1884.6	60.7	3128	9	AB030036	AB030036 Homo sapi
	11	1883.2	60.6	3147	6	AR081724	AR081724 Sequence
	12	1883.2	60.6	3147	6	AR229704	AR229704 Sequence
c	13	1883.2	60.6	3147	6	AR229712	AR229712 Sequence
	14	1883.2	60.6	3147	6	AR430930	AR430930 Sequence
c	15	1883.2	60.6	3147	6	AR430938	AR430938 Sequence
	16	1883.2	60.6	3147	6	AX207897	AX207897 Sequence
	17	1883.2	60.6	3147	6	AX207945	AX207945 Sequence
	18	1883.2	60.6	3147	9	AF057145	AF057145 Homo sapi
	19	1752.6	56.4	2955	6	BD274670	BD274670 Matriptas
	20	1530.8	49.3	2900	6	AR081725	AR081725 Sequence
	21	1530.8	49.3	2900	6	AR229705	AR229705 Sequence
	22	1530.8	49.3	2900	6	AR430931	AR430931 Sequence
	23	1530.8	49.3	2900	9	HSU20428	U20428 Human SNC19
	24	1518.2	48.9	2838	6	AX921703	AX921703 Sequence
	25	1175.8	37.9	2139	9	BC005826	BC005826 Homo sapi
	26	945.4	30.4	1823	9	BC018146	BC018146 Homo sapi
	27	943.8	30.4	2152	6	AR379612	AR379612 Sequence
	28	809.6	26.1	3487	5	AB038498	AB038498 Xenopus l
	29	701.2	22.6	1553	6	AR263832	AR263832 Sequence
c	30	623.2	20.1	235924	10	AC114542	AC114542 Mus muscu
	31	556.4	17.9	726	6	AX473052	AX473052 Sequence
	32	551.8	17.8	723	6	E13204	E13204 Human cDNA
c	33	359.6	11.6	236387	2	AC096042	AC096042 Rattus no
	34	359.6	11.6	261906	2	AC112632	AC112632 Rattus no
	35	359.6	11.6	277797	2	AC128347	AC128347 Rattus no
	36	306.8	9.9	434	6	AR273737	AR273737 Sequence
	37	306.8	9.9	434	6	AR277318	AR277318 Sequence
	38	306.8	9.9	434	6	AR407593	AR407593 Sequence
	39	306.8	9.9	434	6	AX368770	AX368770 Sequence
	40	265	8.5	591	6	AX261852	AX261852 Sequence
	41	244.8	7.9	3142	10	BC057674	BC057674 Mus muscu
	42	244.8	7.9	3183	10	AY240929	AY240929 Mus muscu
	43	244.8	7.9	3183	10	BC029645	BC029645 Mus muscu
	44	243.6	7.8	1485	6	AX814628	AX814628 Sequence
	45	242.4	7.8	3194	9	AY055384	AY055384 Homo sapi

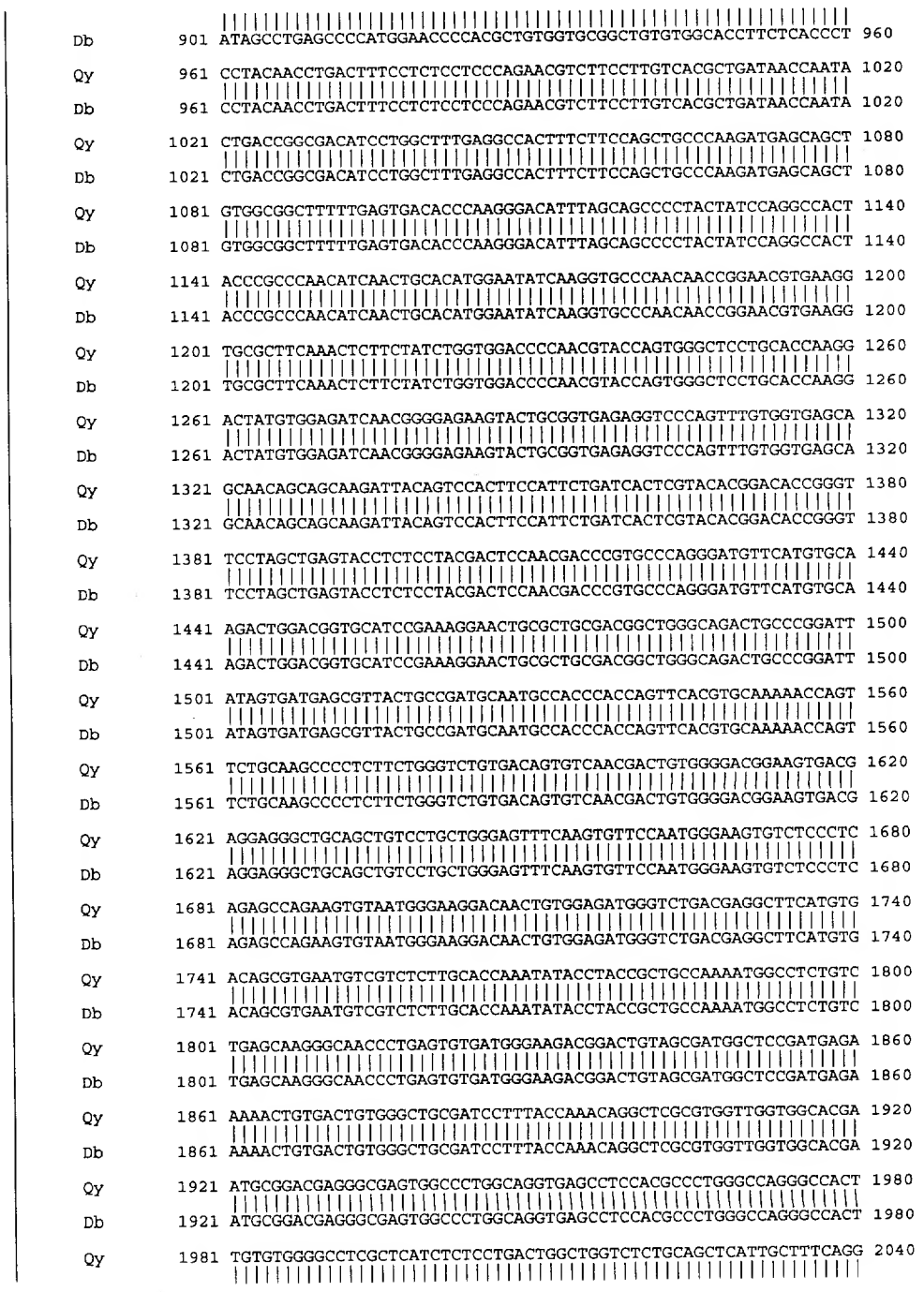
ALIGNMENTS

RESULT 1
AX395268
LOCUS AX395268 3106 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 5 from Patent WO0203787.
ACCESSION AX395268
VERSION AX395268.1 GI:21066293
KEYWORDS .
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Allen,K.D. and Leviten,M.W.
TITLE Transgenic mice containing targeted gene disruptions
JOURNAL Patent: WO 0203787-A 5 17-JAN-2002;

Deltagen, Inc. (US)
FEATURES Location/Qualifiers
source 1. .3106
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN

Query Match 100.0%; Score 3106; DB 6; Length 3106;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





```
Db       1981 TGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGG 2040
Qy       2041 ATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCTGGGTCTGCTGG 2100
Db       2041 ATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCTGGGTCTGCTGG 2100
Qy       2101 ACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACGTATCATCACCC 2160
Db       2101 ACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACGTATCATCACCC 2160
Qy       2161 ACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGAGCTGGAGAAGT 2220
Db       2161 ACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGAGCTGGAGAAGT 2220
Qy       2221 CGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTACCCATGTCTTCC 2280
Db       2221 CGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTACCCATGTCTTCC 2280
Qy       2281 CTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGGAGGTACCGGAG 2340
Db       2281 CTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGGAGGTACCGGAG 2340
Qy       2341 CGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTGTGAGGACCTCA 2400
Db       2341 CGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTGTGAGGACCTCA 2400
Qy       2401 TGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACT 2460
Db       2401 TGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACT 2460
Qy       2461 CCTGCCAGGGTGACTCTGGTGGCCCCCTTGTCAAGCGCGGAGAAAGATGGGCGAATGTTCC 2520
Db       2461 CCTGCCAGGGTGACTCTGGTGGCCCCCTTGTCAAGCGCGGAGAAAGATGGGCGAATGTTCC 2520
Qy       2521 AGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACA 2580
Db       2521 AGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACA 2580
Qy       2581 CAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATAGCAGCATGGAC 2640
Db       2581 CAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATAGCAGCATGGAC 2640
Qy       2641 AGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGGATACAGGAGAG 2700
Db       2641 AGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGGATACAGGAGAG 2700
Qy       2701 GAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACAACCCAGACTGTGAACT 2760
Db       2701 GAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACAACCCAGACTGTGAACT 2760
Qy       2761 GCATCCTTAGGACTCAGAGTTCTTCCAAAGTGGGACCCCCTCAAGAGTTGGAGAGAGAAC 2820
Db       2761 GCATCCTTAGGACTCAGAGTTCTTCCAAAGTGGGACCCCCTCAAGAGTTGGAGAGAGAAC 2820
Qy       2821 TTGCGTGCTAGCGGCCCAGCCTGGGGGCAAGGGTTTGATGGCAGCCTTCCCCCTCTAGCC 2880
Db       2821 TTGCGTGCTAGCGGCCCAGCCTGGGGGCAAGGGTTTGATGGCAGCCTTCCCCCTCTAGCC 2880
Qy       2881 CTGAGCTGGGTGAAGATGATGCTGTCCCGGAGAGCTGCTTCCAACTGTCATTGAGCTCCC 2940
Db       2881 CTGAGCTGGGTGAAGATGATGCTGTCCCGGAGAGCTGCTTCCAACTGTCATTGAGCTCCC 2940
Qy       2941 GGGAGCCCTATGGGAGGAGGGGCTCAGGGTCACTCTTTTCAGGAAGCGCCAGCCCTAGGA 3000
Db       2941 GGGAGCCCTATGGGAGGAGGGGCTCAGGGTCACTCTTTTCAGGAAGCGCCAGCCCTAGGA 3000
Qy       3001 ACCCCAGAAAAGAGTGGTACCTAAGGCTGAAATTGTTTTGCTGTTGCCAGGGGTGGGTAT 3060
Db       3001 ACCCCAGAAAAGAGTGGTACCTAAGGCTGAAATTGTTTTGCTGTTGCCAGGGGTGGGTAT 3060
Qy       3061 TTGAGAGTAAAACATTTTATTTCTTTTTAAAAAAAAAAAAAAAAAA 3106
Db       3061 TTGAGAGTAAAACATTTTATTTCTTTTTAAAAAAAAAAAAAAAAAA 3106
```

```
RESULT 2
AF042822
LOCUS       AF042822               3106 bp    mRNA    linear   ROD 27-MAR-2000
DEFINITION  Mus musculus epithin mRNA, complete cds.
ACCESSION   AF042822
VERSION     AF042822.3  GI:7330637
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3106)
  AUTHORS   Kim,M.G., Chen,C., Lyu,M.S., Cho,E.G., Park,D., Kozak,C. and
            Schwartz,R.H.
  TITLE     Cloning and chromosomal mapping of a gene isolated from thymic
            stromal cells encoding a new mouse type II membrane serine
            protease, epithin, containing four LDL receptor modules and two CUB
            domains
  JOURNAL   Immunogenetics 49 (5), 420-428 (1999)
  MEDLINE   99216440
   PUBMED   10199918
REFERENCE   2  (bases 1 to 3106)
  AUTHORS   Kim,M.G., Chen,C., Cho,E., Park,D. and Schwartz,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-1998) Laboratory of Cellular and Molecular
            Immunology, NIAID, NIH, 9000 Rockville Pike, Bethesda, MD 20892,
            USA
REFERENCE   3  (bases 1 to 3106)
  AUTHORS   Kim,M.G., Chen,C., Cho,E., Park,D. and Schwartz,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-1999) Laboratory of Cellular and Molecular
            Immunology, NIAID, NIH, 9000 Rockville Pike, Bethesda, MD 20892,
            USA
  REMARK    Sequence update by submitter
REFERENCE   4  (bases 1 to 3106)
  AUTHORS   Kim,M.G., Chen,C., Cho,E., Park,D. and Schwartz,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2000) Laboratory of Cellular and Molecular
            Immunology, NIAID, NIH, 9000 Rockville Pike, Bethesda, MD 20892,
            USA
  REMARK    Sequence update by submitter
COMMENT     On Mar 27, 2000 this sequence version replaced gi:6435480.
FEATURES             Location/Qualifiers
     source          1. .3106
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C.B.17scid"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /cell_type="stromal"
                     /tissue_type="thymus"
                     /note="severe combined immunodeficient (SCID)"
     CDS             63. .2630
                     /function="serine protease"
                     /codon_start=1
                     /product="epithin"
                     /protein_id="AAD02230.3"
                     /db_xref="GI:7330638"
                     /translation="MGSNRGRKAGGGSQDFGAGLKYNSRLENMNGFEEGVEFLPANNA
                     KKVEKRGPRRWVVLVAVLFSFLLLSLMAGLLVWHFHYRNVRVQKVFNGHLRITNEIFL
                     DAYENSTSTEFISLASQVKEALKLLYNEVPVLGPYHKKSAVTAFSEGSVIAYYWSEFS
                     IPPHLAEEVDRAMAVERVVTLPPRARALKSFVLTSVVAFPIDPRMLQRTQDNSCSFAL
                     HAHGAAVTRFTTPGFPNSPYPAHARCQWVLRGDADSVLSLTFRSFDVAPCDEHGSDLV
                     TVYDSLSPMEPHAVVRLCGTFSPSYNLTFLSSQNVFLVTLITNTDRRHPGFEATFFQL
                     PKMSSCGGFLSDTQGTFSSPYYPGHYPPNINCTWNIKVPNNRNVKVRFKLFYLVDPNV
                     PVGSCTKDYVEINGEKYCGERSQFVVSSNSSKITVHFHSDHSYTDTGFLAEYLSYDSN
                     DPCPGMFMCKTGRCIRKELRCDGWADCPDYSDERYCRCNATHQFTCKNQFCKPLFWVC
                     DSVNDCGDGSDEEGCSCPAGSFKCSNGKCLPQSQKCNGKDNCGDGSDEASCDSVNVVS
                     CTKYTYRCQNGLCLSKGNPECDGKTDCSDGSDEKNCDCGLRSFTKQARVVGGTNADEG
                     EWPWQVSLHALGQGHLCGASLISPDWLVSAAHCFQDDKNFKYSDYTMWTAFLGLLDQS
                     KRSASGVQELKLKRIITHPSFNDFTFDYDIALLELEKSVEYSTVVRPICLPDATHVFP
```

AGKAIWVTGWGHTKEGGTGALILQKGEIRVINQTTCEDLMPQQITPRMMCVGFLSGGV
DSCQGDSGGPLSSAEKDGRMFQAGVVSWGEGCAQRNKPGVYTRLPVVRDWIKEHTGV"

ORIGIN

Query Match 100.0%; Score 3106; DB 10; Length 3106;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CATGGTAGACGGCTGCCCGGAGGGACCACGCGTCTGAGACCGGCGATCGGACCGCCAAAA 60
Db       1 CATGGTAGACGGCTGCCCGGAGGGACCACGCGTCTGAGACCGGCGATCGGACCGCCAAAA 60

Qy      61 CCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTTCGGCGCGGGAC 120
Db      61 CCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTTCGGCGCGGGAC 120

Qy     121 TCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGC 180
Db     121 TCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGC 180

Qy     181 CTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGCTGGTGG 240
Db     181 CTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGCTGGTGG 240

Qy     241 CAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGTGTGGCACTTCC 300
Db     241 CAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGTGTGGCACTTCC 300

Qy     301 ATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAAATGAGA 360
Db     301 ATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAAATGAGA 360

Qy     361 TCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCTGGCCAGCCAGG 420
Db     361 TCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCTGGCCAGCCAGG 420

Qy     421 TGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCCCTACCACAAGA 480
Db     421 TGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCCCTACCACAAGA 480

Qy     481 AGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTGGTCAGAGTTCA 540
Db     481 AGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTGGTCAGAGTTCA 540

Qy     541 GCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGAGCGAGTTGTAA 600
Db     541 GCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGAGCGAGTTGTAA 600

Qy     601 CATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGTGGCCTTCC 660
Db     601 CATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGTGGCCTTCC 660

Qy     661 CCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTGCATG 720
Db     661 CCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTGCATG 720

Qy     721 CCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAGTCCCTACCCGG 780
Db     721 CCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAGTCCCTACCCGG 780

Qy     781 CGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCTGTGCTGAGCCTCACCT 840
Db     781 CGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCTGTGCTGAGCCTCACCT 840

Qy     841 TCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGTCACCGTGTATG 900
Db     841 TCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGTCACCGTGTATG 900

Qy     901 ATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCACCTTCTCACCCT 960
Db     901 ATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCACCTTCTCACCCT 960

Qy     961 CCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCTGATAACCAATA 1020
Db     961 CCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCTGATAACCAATA 1020

Qy    1021 CTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCT 1080
Db    1021 CTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCT 1080

Qy    1081 GTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTATCCAGGCCACT 1140
Db    1081 GTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTATCCAGGCCACT 1140

Qy    1141 ACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAGG 1200
Db    1141 ACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAGG 1200

Qy    1201 TGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTCCTGCACCAAGG 1260
Db    1201 TGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTCCTGCACCAAGG 1260

Qy    1261 ACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCA 1320
Db    1261 ACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCA 1320

Qy    1321 GCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCGTACACGGACACCGGGT 1380
Db    1321 GCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCGTACACGGACACCGGGT 1380

Qy    1381 TCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCAGGGATGTTCATGTGCA 1440
Db    1381 TCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCAGGGATGTTCATGTGCA 1440

Qy    1441 AGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGGGCAGACTGCCCGGATT 1500
Db    1441 AGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGGGCAGACTGCCCGGATT 1500

Qy    1501 ATAGTGATGAGCGTTACTGCCGATGCAATGCCCACCCACCAGTTCACGTGCAAAAACCAGT 1560
Db    1501 ATAGTGATGAGCGTTACTGCCGATGCAATGCCCACCCACCAGTTCACGTGCAAAAACCAGT 1560

Qy    1561 TCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGACGGAAGTGACG 1620
Db    1561 TCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGACGGAAGTGACG 1620

Qy    1621 AGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAATGGGAAGTGTCTCCCTC 1680
Db    1621 AGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAATGGGAAGTGTCTCCCTC 1680

Qy    1681 AGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCTGACGAGGCTTCATGTG 1740
Db    1681 AGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCTGACGAGGCTTCATGTG 1740

Qy    1741 ACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAATGGCCTCTGTC 1800
Db    1741 ACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAATGGCCTCTGTC 1800

Qy    1801 TGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGCGATGGCTCCGATGAGA 1860
Db    1801 TGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGCGATGGCTCCGATGAGA 1860

Qy    1861 AAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGCGTGGTTGGTGGCACGA 1920
Db    1861 AAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGCGTGGTTGGTGGCACGA 1920

Qy    1921 ATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCCCTGGGCCAGGGCCACT 1980
Db    1921 ATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCCCTGGGCCAGGGCCACT 1980

Qy    1981 TGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGG 2040
Db    1981 TGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGG 2040

Qy    2041 ATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCTGGGTCTGCTGG 2100
```

```
Db       2041 ATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCTGGGTCTGCTGG 2100
Qy       2101 ACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACGTATCATCACCC 2160
Db       2101 ACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACGTATCATCACCC 2160
Qy       2161 ACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGAGCTGGAGAAGT 2220
Db       2161 ACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGAGCTGGAGAAGT 2220
Qy       2221 CGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTACCCATGTCTTCC 2280
Db       2221 CGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTACCCATGTCTTCC 2280
Qy       2281 CTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGGAGGTACCGGAG 2340
Db       2281 CTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGGAGGTACCGGAG 2340
Qy       2341 CGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTGTGAGGACCTCA 2400
Db       2341 CGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTGTGAGGACCTCA 2400
Qy       2401 TGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACT 2460
Db       2401 TGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACT 2460
Qy       2461 CCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAAGATGGGCGAATGTTCC 2520
Db       2461 CCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAAGATGGGCGAATGTTCC 2520
Qy       2521 AGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACA 2580
Db       2521 AGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACA 2580
Qy       2581 CAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATAGCAGCATGGAC 2640
Db       2581 CAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATAGCAGCATGGAC 2640
Qy       2641 AGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGGATACAGGAGAG 2700
Db       2641 AGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGGATACAGGAGAG 2700
Qy       2701 GAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCCAACACAACCCAGACTGTGAACT 2760
Db       2701 GAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCCAACACAACCCAGACTGTGAACT 2760
Qy       2761 GCATCCTTAGGACTCAGAGTTCTTCCAAAGTGGGACCCCCTCAAGAGTTGGAGAGAGAAC 2820
Db       2761 GCATCCTTAGGACTCAGAGTTCTTCCAAAGTGGGACCCCCTCAAGAGTTGGAGAGAGAAC 2820
Qy       2821 TTGCGTGCTAGCGGCCCAGCCTGGGGGCAAGGGTTTGATGGCAGCCTTCCCCCTCTAGCC 2880
Db       2821 TTGCGTGCTAGCGGCCCAGCCTGGGGGCAAGGGTTTGATGGCAGCCTTCCCCCTCTAGCC 2880
Qy       2881 CTGAGCTGGGTGAAGATGATGCTGTCCCGGAGAGCTGCTTCCAACTGTCATTGAGCTCCC 2940
Db       2881 CTGAGCTGGGTGAAGATGATGCTGTCCCGGAGAGCTGCTTCCAACTGTCATTGAGCTCCC 2940
Qy       2941 GGGAGCCCTATGGGAGGAGGGGCTCAGGGTCACTCTTTTCAGGAAGCGCCAGCCCTAGGA 3000
Db       2941 GGGAGCCCTATGGGAGGAGGGGCTCAGGGTCACTCTTTTCAGGAAGCGCCAGCCCTAGGA 3000
Qy       3001 ACCCCAGAAAAGAGTGGTACCTAAGGCTGAAATTGTTTTGCTGTTGCCAGGGGTGGGTAT 3060
Db       3001 ACCCCAGAAAAGAGTGGTACCTAAGGCTGAAATTGTTTTGCTGTTGCCAGGGGTGGGTAT 3060
Qy       3061 TTGAGAGTAAAACATTTTATTTCTTTTTAAAAAAAAAAAAAAAAAA 3106
Db       3061 TTGAGAGTAAAACATTTTATTTCTTTTTAAAAAAAAAAAAAAAAAA 3106
```

RESULT 3
BC005496

```
LOCUS       BC005496               3248 bp    mRNA    linear   ROD 04-NOV-2003
DEFINITION  Mus musculus suppression of tumorigenicity 14 (colon carcinoma),
            mRNA (cDNA clone MGC:7395 IMAGE:3488059), complete cds.
ACCESSION   BC005496
VERSION     BC005496.1  GI:13529565
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3248)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 3248)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-MAR-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 7 Row: i Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1. .3248
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="MGC:7395 IMAGE:3488059"
                     /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
                     old, gross tissue."
                     /clone_lib="NCI_CGAP_Mam5"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .3248
```

```
                     /gene="St14"
                     /note="synonym: Epithin"
                     /db_xref="LocusID:19143"
                     /db_xref="MGI:1338881"
     CDS             173. .2740
                     /codon_start=1
                     /product="suppression of tumorigenicity 14 (colon
                     carcinoma)"
                     /protein_id="AAH05496.1"
                     /db_xref="GI:13529566"
                     /db_xref="LocusID:19143"
                     /translation="MGSNRGRKAGGGSQDFGAGLKYNSRLENMNGFEEGVEFLPANNA
                     KKVEKRGPRRWVVLVAVLFSFLLLSLMAGLLVWHFHYRNVRVQKVFNGHLRITNEIFL
                     DAYENSTSTEFISLASQVKEALKLLYNEVPVLGPYHKKSAVTAFSEGSVIAYYWSEFS
                     IPPHLAEEVDRAMAVERVVTLPPRARALKSFVLTSVVAFPIDPRMLQRTQDNSCSFAL
                     HAHGAAVTRFTTPGFPNSPYPAHARCQWVLRGDADSVLSLTFRSFDVAPCDEHGSDLV
                     TVYDSLSPMEPHAVVRLCGTFSPSYNLTFLSSQNVFLVTLITNTDRRHPGFEATFFQL
                     PKMSSCGGFLSDTQGTFSSPYYPGHYPPNINCTWNIKVPNNRNVKVRFKLFYLVDPNV
                     PVGSCTKDYVEINGEKYCGERSQFVVSSNSSKITVHFHSDHSYTDTGFLAEYLSYDSN
                     DPCPGMFMCKTGRCIRKELRCDGWADCPDYSDERYCRCNATHQFTCKNQFCKPLFWVC
                     DSVNDCGDGSDEEGCSCPAGSFKCSNGKCLPQSQKCNGKDNCGDGSDEASCDSVNVVS
                     CTKYTYRCQNGLCLSKGNPECDGKTDCSDGSDEKNCDCGLRSFTKQARVVGGTNADEG
                     EWPWQVSLHALGQGHLCGASLISPDWLVSAAHCFQDDKNFKYSDYTMWTAFLGLLDQS
                     KRSASGVQELKLKRIITHPSFNDFTFDYDIALLELEKSVEYSTVVRPICLPDATHVFP
                     AGKAIWVTGWGHTKEGGTGALILQKGEIRVINQTTCEDLMPQQITPRMMCVGFLSGGV
                     DSCQGDSGGPLSSAEKDGRMFQAGVVSWGEGCAQRNKPGVYTRLPVVRDWIKEHTGV"
     misc_feature    833. .1165
                     /note="CUB; Region: CUB domain"
                     /db_xref="CDD:cd00041"
     misc_feature    1190. .1504
                     /note="CUB; Region: CUB domain"
                     /db_xref="CDD:cd00041"
     misc_feature    1532. .1636
                     /note="LDLa; Region: Low-density lipoprotein receptor
                     domain class A"
                     /db_xref="CDD:smart00192"
     misc_feature    1640. .1747
                     /note="LDLa; Region: Low-density lipoprotein receptor
                     domain class A"
                     /db_xref="CDD:smart00192"
     misc_feature    1745. .1855
                     /note="LDLa; Region: Low-density lipoprotein receptor
                     domain class A"
                     /db_xref="CDD:smart00192"
     misc_feature    1871. .1978
                     /note="LDLa; Region: Low Density Lipoprotein Receptor
                     Class A domain, a cysteine-rich repeat that plays a
                     central role in mammalian cholesterol metabolism"
                     /db_xref="CDD:cd00112"
     misc_feature    2015. .2728
                     /note="Tryp_SPc; Region: Trypsin-like serine protease"
                     /db_xref="CDD:cd00190"
ORIGIN

  Query Match             99.9%;  Score 3102.8;  DB 10;  Length 3248;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3104;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CATGGTAGACGGCTGCCCGGAGGGACCACGCGTCTGAGACCGGCGATCGGACCGCCAAAA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 CATGGTAGACGGCTGCCCGGAGGGACCACGCGTCTGAGACCGGCGATCGGACCGCCAAAA 170

Qy         61 CCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTTCGGCGCGGGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 CCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTTCGGCGCGGGAC 230

Qy        121 TCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGC 290

Qy        181 CTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGCTGGTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 CTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGCTGGTGG 350

Qy        241 CAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGTGTGGCACTTCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 CAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGTGTGGCACTTCC 410

Qy        301 ATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAAATGAGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 ATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAAATGAGA 470

Qy        361 TCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCTGGCCAGCCAGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 TCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCTGGCCAGCCAGG 530

Qy        421 TGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCCCTACCACAAGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 TGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCCCTACCACAAGA 590

Qy        481 AGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTGGTCAGAGTTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 AGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTGGTCAGAGTTCA 650

Qy        541 GCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGAGCGAGTTGTAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 GCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGAGCGAGTTGTAA 710

Qy        601 CATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGTGGCCTTCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 CATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGTGGCCTTCC 770

Qy        661 CCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTGCATG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 CCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTGCATG 830

Qy        721 CCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAGTCCCTACCCGG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        831 CCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAGTCCCTACCCGG 890

Qy        781 CGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGGACGCCGACTCTGTGCTGAGCCTCACCT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        891 CGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGGACGCCGACTCTGTGCTGAGCCTCACCT 950

Qy        841 TCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGTCACCGTGTATG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        951 TCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGTCACCGTGTATG 1010

Qy        901 ATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCACCTTCTCACCCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1011 ATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCACCTTCTCACCCT 1070

Qy        961 CCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCTGATAACCAATA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1071 CCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCTGATAACCAATA 1130

Qy       1021 CTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1131 CTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCT 1190

Qy       1081 GTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTATCCAGGCCACT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1191 GTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTATCCAGGCCACT 1250

Qy       1141 ACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1251 ACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAGG 1310

Qy       1201 TGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTCCTGCACCAAGG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1311 TGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTCCTGCACCAAGG 1370

Qy       1261 ACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1371 ACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCA 1430
```

```
Qy       1321 GCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCGTACACGGACACCGGGT 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1431 GCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCGTACACGGACACCGGGT 1490

Qy       1381 TCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCAGGGATGTTCATGTGCA 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1491 TCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCAGGGATGTTCATGTGCA 1550

Qy       1441 AGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGGGCAGACTGCCCGGATT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1551 AGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGGGCAGACTGCCCGGATT 1610

Qy       1501 ATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTCACGTGCAAAAACCAGT 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1611 ATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTCACGTGCAAAAACCAGT 1670

Qy       1561 TCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGACGGAAGTGACG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1671 TCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGACGGAAGTGACG 1730

Qy       1621 AGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAATGGGAAGTGTCTCCCTC 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1731 AGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAATGGGAAGTGTCTCCCTC 1790

Qy       1681 AGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCTGACGAGGCTTCATGTG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1791 AGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCTGACGAGGCTTCATGTG 1850

Qy       1741 ACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAATGGCCTCTGTC 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1851 ACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAATGGCCTCTGTC 1910

Qy       1801 TGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGCGATGGCTCCGATGAGA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1911 TGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGCGATGGCTCCGATGAGA 1970

Qy       1861 AAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGCGTGGTTGGTGGCACGA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1971 AAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGCGTGGTTGGTGGCACGA 2030

Qy       1921 ATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCCCTGGGCCAGGGCCACT 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2031 ATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCCCTGGGCCAGGGCCACT 2090

Qy       1981 TGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2091 TGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGG 2150

Qy       2041 ATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCTGGGTCTGCTGG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2151 ATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCTGGGTCTGCTGG 2210

Qy       2101 ACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACGTATCATCACCC 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2211 ACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACGTATCATCACCC 2270

Qy       2161 ACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGAGCTGGAGAAGT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2271 ACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGAGCTGGAGAAGT 2330

Qy       2221 CGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTACCCATGTCTTCC 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2331 CGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTACCCATGTCTTCC 2390

Qy       2281 CTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGGAGGTACCGGAG 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2391 CTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGGAGGTACCGGAG 2450

Qy       2341 CGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTGTGAGGACCTCA 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2451 CGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTGTGAGGACCTCA 2510

Qy       2401 TGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2511 TGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACT 2570

Qy       2461 CCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAAGATGGGCGAATGTTCC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2571 CCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAAGATGGGCGAATGTTCC 2630

Qy       2521 AGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACA 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2631 AGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACA 2690

Qy       2581 CAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATAGCAGCATGGAC 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2691 CAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATAGCAGCATGGAC 2750

Qy       2641 AGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGGATACAGGAGAG 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2751 AGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGGATACAGGAGAG 2810

Qy       2701 GAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACAACCCAGACTGTGAACT 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2811 GAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACAACCCAGACTGTGAACT 2870

Qy       2761 GCATCCTTAGGACTCAGAGTTCTTCCAAAGTGGGACCCCCTCAAGAGTTGGAGAGAGAAC 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2871 GCATCCTTAGGACTCAGAGTTCTTCCAAAGTGGGACCCCCTCAAGAGTTGGAGAGAGAAC 2930

Qy       2821 TTGCGTGCTAGCGGCCCAGCCTGGGGGCAAGGGTTTGATGGCAGCCTTCCCCCTCTAGCC 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2931 TTGCGTGCTAGCGGCCCAGCCTGGGGGCAAGGGTTTGATGGCAGCCTTCCCCCTCTAGCC 2990

Qy       2881 CTGAGCTGGGTGAAGATGATGCTGTCCCGGAGAGCTGCTTCCAACTGTCATTGAGCTCCC 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2991 CTGAGCTGGGTGAAGATGATGCTGTCCCGGAGAGCTGCTTCCAACTGTCATTGAGCTCCC 3050

Qy       2941 GGGAGCCCTATGGGAGGAGGGGCTCAGGGTCACTCTTTTCAGGAAGCGCCAGCCCTAGGA 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3051 GGGAGCCCTATGGGAGGAGGGGCTCAGGGTCACTCTTTTCAGGAAGCGCCAGCCCTAGGA 3110

Qy       3001 ACCCCAGAAAAGAGTGGTACCTAAGGCTGAAATTGTTTTGCTGTTGCCAGGGGTGGGTAT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3111 ACCCCAGAAAAGAGTGGTACCTAAGGCTGAAATTGTTTTGCTGTTGCCAGGGGTGGGTAT 3170

Qy       3061 TTGAGAGTAAAACATTTTATTTCTTTTTAAAAAAAAAAAAAAAAAA 3106
             ||||||||||||||||||||||||||  ||||||||||||||||||
Db       3171 TTGAGAGTAAAACATTTTATTTCTTTAAAAAAAAAAAAAAAAAAA 3216
```

RESULT 4
AB049189

```
LOCUS       AB049189                3174 bp    mRNA    linear   ROD 14-APR-2001
DEFINITION  Rattus norvegicus mRNA for membrane bound arginine specific serine
            protease, complete cds.
ACCESSION   AB049189
VERSION     AB049189.1  GI:10336526
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (sites)
  AUTHORS   Inoue,H., Takahashi,K. and Kishi,K.
  TITLE     membrane-bound arginine specific serine protease
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 3174)
  AUTHORS   Inoue,H., Takahashi,K. and Kishi,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-2000) Hideshi Inoue, Tokyo University of Pharmacy
            and Life Science, School of Life Science; 1432-1 Horinouchi,
            Hachioji-shi, Tokyo 192-0392, Japan (E-mail:hinoue@ls.toyaku.ac.jp,
            Tel:81-426-76-7153, Fax:81-426-76-7157)
FEATURES             Location/Qualifiers
```

```
     source          1. .3174
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /tissue_type="duodenum"
                     /dev_stage="7 weeks old"
     CDS             127. .2694
                     /codon_start=1
                     /evidence=experimental
                     /product="membrane bound arginine specific serine
                     protease"
                     /protein_id="BAB13765.1"
                     /db_xref="GI:10336527"
                     /translation="MGNNRGRKAGGGSQDFGAGLKYNSRLENMNGFEEGVEFLPVNNA
                     KQVEKRGPRRWVVMVAVVFSFLLLSLMAGLLVWHFHYRNVRIQKVFNGHLRITNENFL
                     DAYENSTSTEFISLASQVKEALKLMYSEVPVLGPYHKKSTVTAFSEGSVIAYYWSEFS
                     IPPHLEEEVDRAMAVERVVTLPPRARALKSFVLTSVVAFPIDPRMLQRTQDNSCSFAL
                     HARGRTVTRFTTPGFPNSPYPAHARCQWVLRGDADSVLSLTFRSFDVAPCDGHDSDLV
                     TVYDSLSPMEPHAVVRLCGTFSPSYNLTFLSSQNVFLVTLITNTDRRHPGFEATFFQL
                     PKMSSCGGLLSEAQGTFSSPYYPGHYPPNINCTWNIKVPNNRNVKVRFKLFYLVDPNI
                     PVGSCTKDYVEINGEKFCGERSQFVVSSNSSKITVHFHSDHSYTDTGFLAEYLSYDSN
                     DPCPGMFMCKTGRCIRKDLRCDGWADCPDYSDERHCRCNATHQFMCKNQFCKPLFWVC
                     DSVNDCGDGSDEEGCSCPAGSFKCSNGKCLPQSQQCNGKDDCGDGSDEASCDNVNAVS
                     CTKYTYRCQNGLCLNKGNPECDGKKDCSDGSDEKNCDCGLRSFTKQARVVGGTNADEG
                     EWPWQVSLHALGQGHLCGASLISPDWLVSAAHCFQDETIFKYSDHTMWTAFLGLLDQS
                     KRSASGVQEHKLKRIITHPSFNDFTFDYDIALLELEKPAEYSTVVRPICLPDNTHVFP
                     AGKAIWVTGWGHTKEGGTGALILQKGEIRVINQTTCEELLPQQITPRMMCVGFLSGGV
                     DSCQGDSGGPLSSVEKDGRIFQAGVVSWGEGCAQRNKPGVYTRIPEVRDWIKEQTGV"
     variation       360
                     /note="compared with the sequence reported in Acc#
                     AB037898"
                     /replace="t"
     variation       513
                     /note="compared with the sequence reported in Acc#
                     AB037898"
                     /replace="t"
     variation       573
                     /note="compared with the sequence reported in Acc#
                     AB037898"
                     /replace="t"
     variation       2121
                     /note="compared with the sequence reported in Acc#
                     AB037898"
                     /replace="t"
     polyA_site      3157. .3174
ORIGIN

  Query Match             83.5%;  Score 2595;  DB 10;  Length 3174;
  Best Local Similarity   91.1%;  Pred. No. 0;
  Matches 2853;  Conservative    0;  Mismatches  230;  Indels   50;  Gaps    7;

Qy          1 CATGGTAGACGGCTGCCCGGAGGGACCACGCGTCTGAGACCGGCGATCGGACCGCCAAAA 60
              |||||||| ||| |||| | || ||||||||||||||||||||||||||| ||||| | |
Db         65 CATGGTAGTCGGTTGCCGGCAGAGACCACGCGTCTGAGACCGGCGATCGGGCCGCCGAGA 124

Qy         61 CCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTTCGGCGCGGGAC 120
              ||||||| | |||||||||||||||||||||||||||||| |||||||||||||| ||||
Db        125 CCATGGGGAACAATCGGGGCCGCAAGGCCGGAGGGGGCTCGCAGGACTTCGGCGCCGGAC 184

Qy        121 TCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGC 244

Qy        181 CTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGCTGGTGG 240
              ||| |||||| |||||| | ||||||||||||||||||||||||||||||| ||||||
Db        245 CTGTGAACAACGCCAAGCAGGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGATGGTGG 304

Qy        241 CAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGTGTGGCACTTCC 300
              | ||  |||||||||||| |||||||| ||||||||||||||||||||||||||||||||
Db        305 CCGTAGTGTTCAGCTTCCTTTTGCTCTCGCTCATGGCCGGCTTGCTGGTGTGGCACTTCC 364

Qy        301 ATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAAATGAGA 360
              |||| |||||||||| || | ||||| ||||||||||||||||||||||||| |||||||
Db        365 ATTACCGGAATGTGAGGATCCAAAAGGTCTTCAATGGCCATCTGAGGATCACGAATGAGA 424

Qy        361 TCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCTGGCCAGCCAGG 420
              |||||||| |||||||||||||||||||||||| ||||| |||||||||||||| ||||
Db        425 ACTTTCTGGACGCGTATGAGAACTCCACCTCCACGGAGTTCATCAGCCTGGCCAGTCAGG 484

Qy        421 TGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCCCTACCACAAGA 480
              ||||||| |||||||||||| |||||| |||||||||||||||||| ||||||||||||
Db        485 TGAAGGAAGCGCTGAAGCTGATGTACAGTGAAGTCCCTGTCCTGGGCCCCTACCACAAGA 544

Qy        481 AGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTGGTCAGAGTTCA 540
              | ||| |||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Db        545 AATCGACTGTAACTGCCTTCAGTGAGGGCAGCGTCATTGCCTACTACTGGTCAGAGTTCA 604

Qy        541 GCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGAGCGAGTTGTAA 600
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |
Db        605 GCATCCCCCCACACCTGGAGGAAGAGGTTGATCGCGCCATGGCTGTGGAGCGAGTTGTCA 664

Qy        601 CATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGTGGCCTTCC 660
              || |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db        665 CACTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTGACCTCTGTGGTGGCCTTCC 724

Qy        661 CCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTGCATG 720
              |||||||||||||||||| |||||||||||||||||||||||||||||||| ||||
Db        725 CCATTGACCCCAGAATGCTCCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTCCATG 784

Qy        721 CCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAGTCCCTACCCGG 780
              ||| |||      |||||||||| |||||||||||||||||||| |||||||||||||| |
Db        785 CCCGTGGCAGGACAGTGACACGTTTCACTACCCCTGGCTTCCCTAACAGTCCCTACCCAG 844

Qy        781 CGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCTGTGCTGAGCCTCACCT 840
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        845 CTCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCGGACTCTGTGCTGAGCCTCACCT 904

Qy        841 TCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGTCACCGTGTATG 900
              |||||| |||||||||| |||||||||| ||||| |||||||||||||| |||||||
Db        905 TCCGAAGTTTTGATGTCGCGCCCTGTGATGGGCATGACAGTGACCTGGTCACAGTGTATG 964

Qy        901 ATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCACCTTCTCACCCT 960
              | |||||||||||||||| ||||| ||||||||||||||||||||||||||| ||||
Db        965 ACAGCCTGAGCCCCATGGAGCCCCATGCTGTGGTGCGGCTGTGTGGCACCTTCTCGCCCT 1024

Qy        961 CCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCTGATAACCAATA 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |
Db       1025 CCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTTACGCTGATAACCAACA 1084

Qy       1021 CTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1085 CTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCT 1144

Qy       1081 GTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTATCCAGGCCACT 1140
              |||||||| |||||||||| |||||||||| |||||||||||||||||||||||| |
Db       1145 GTGGCGGCCTTTTGAGTGAAGCCCAAGGGACGTTTAGCAGCCCCTACTATCCAGGCCATT 1204

Qy       1141 ACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAGG 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1205 ACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAAG 1264

Qy       1201 TGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTCCTGCACCAAGG 1260
              | |||||||||||||||||||||||||||||||||||||| |||||||||  |||||||||
Db       1265 TACGCTTCAAACTCTTCTATCTGGTGGACCCCAACATACCAGTGGGAAGCTGCACCAAGG 1324

Qy       1261 ACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCA 1320
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db       1325 ACTATGTGGAGATCAACGGGGAGAAGTTCTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCA 1384

Qy       1321 GCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCGTACACGGACACCGGGT 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1385 GCAACAGCAGCAAGATTACGGTCCACTTCCATTCTGATCACTCGTACACGGACACCGGGT 1444

Qy       1381 TCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCAGGGATGTTCATGTGCA 1440
```

```
Db      1445 TCCTGGCTGAGTATCTCTCCTACGACTCCAATGACCCATGCCCAGGGATGTTCATGTGTA 1504
Qy      1441 AGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGGGCAGACTGCCCGGATT 1500
Db      1505 AGACTGGACGGTGCATCCGCAAGGACTTGCGCTGCGATGGCTGGGCTGACTGCCCGGACT 1564
Qy      1501 ATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTCACGTGCAAAAACCAGT 1560
Db      1565 ACAGTGATGAGCGTCACTGCCGATGCAACGCTACCCACCAGTTCATGTGCAAGAACCAGT 1624
Qy      1561 TCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGACGGAAGTGACG 1620
Db      1625 TCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGATGGCAGCGATG 1684
Qy      1621 AGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAATGGGAAGTGTCTCCCTC 1680
Db      1685 AGGAGGGCTGCAGTTGTCCTGCCGGGAGTTTCAAGTGTTCCAATGGGAAGTGTCTCCCTC 1744
Qy      1681 AGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCTGACGAGGCTTCATGTG 1740
Db      1745 AGAGCCAGCAGTGTAATGGGAAGGACGACTGTGGAGATGGGTCTGACGAGGCCTCATGTG 1804
Qy      1741 ACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAATGGCCTCTGTC 1800
Db      1805 ACAATGTGAATGCCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAATGGCCTCTGTC 1864
Qy      1801 TGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGCGATGGCTCCGATGAGA 1860
Db      1865 TGAACAAGGGCAACCCTGAGTGTGATGGGAAGAAGGACTGCAGCGATGGCTCGGATGAGA 1924
Qy      1861 AAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGCGTGGTTGGTGGCACGA 1920
Db      1925 AAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGAGTGGTTGGTGGCACGA 1984
Qy      1921 ATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCCCTGGGCCAGGGCCACT 1980
Db      1985 ATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCTCTGGGCCAGGGCCACT 2044
Qy      1981 TGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGG 2040
Db      2045 TGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGG 2104
Qy      2041 ATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCTGGGTCTGCTGG 2100
Db      2105 ATGAGACAATTTTCAAGTACTCAGACCACACCATGTGGACAGCCTTCCTCGGTCTGCTGG 2164
Qy      2101 ACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACGTATCATCACCC 2160
Db      2165 ACCAGAGCAAGCGCAGTGCCTCTGGGGTACAGGAGCACAAGCTCAAACGTATCATCACTC 2224
Qy      2161 ACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGAGCTGGAGAAGT 2220
Db      2225 ACCCTTCCTTTAATGACTTTACCTTCGACTATGACATTGCCTTGCTGGAGCTGGAGAAGC 2284
Qy      2221 CGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTACCCATGTCTTCC 2280
Db      2285 CGGCAGAGTACAGCACTGTCGTGCGCCCCATCTGCCTGCCTGATAATACCCATGTCTTTC 2344
Qy      2281 CTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGGAGGTACCGGAG 2340
Db      2345 CTGCCGGCAAAGCCATCTGGGTCACAGGCTGGGGCCACACGAAAGAAGGAGGAACTGGAG 2404
Qy      2341 CGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTGTGAGGACCTCA 2400
Db      2405 CACTGATCCTGCAGAAGGGAGAGATCCGGGTCATCAACCAGACCACCTGTGAGGAACTCT 2464
Qy      2401 TGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACT 2460
Db      2465 TGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACT 2524
Qy      2461 CCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAAGATGGGCGAATGTTCC 2520
Db      2525 CCTGCCAGGGTGATTCCGGTGGCCCCTTGTCCAGTGTGGAGAAAGATGGGCGAATCTTCC 2584
Qy      2521 AGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACA 2580
Db      2585 AGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACA 2644
Qy      2581 CAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATAGCAGCATGGAC 2640
Db      2645 CAAGGATCCCTGAAGTTCGGGACTGGATCAAAGAGCAAACTGGGGTATAGAAGCATGGAC 2704
Qy      2641 AGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGGATACAGGAGAG 2700
Db      2705 AGACAACCCACCATGAACTCGCAAAGGGATGCCCCCCACGCGTACCTGGATACAGCAGAG 2764
Qy      2701 GAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACAACCCAGACTGTGAACT 2760
Db      2765 GATCGCTGAGGACATTTGTGCTGTGGCCTCTCCCC-----CACATCCCAGACTGTGAACT 2819
Qy      2761 GCATCCTTAGGACTCAGAGTTCTTCCA---------AAGTGGGACCCCCTCAAGAGTTGG 2811
Db      2820 GCTTCCCTAGGACTCAGAATCCTTCCCTGAGCTTCCAAAGTGGGACCCCTCAGGAGTTGG 2879
Qy      2812 AGAGAGAACTTGCGTGCTAGCGGCCCAGCCTG----GGGGCAAGGGTTTGATGGCAGCCT 2867
Db      2880 CCAGAGATCTTCTGTGCTGGTGGCCCAGCCCGGGGTGGGGCAAGGATTTGATGGCAGCCT 2939
Qy      2868 TCCCCCTCTAGCCCTGAGCTGGGTGAAGATGATGCTGTCCCGGAGAGCTGCTTCC----- 2922
Db      2940 T-CCCCTCTAACCCTGAACTGGGTAAAGACGATGCTTTCCTGGAGAACTGCTTCCAGACT 2998
Qy      2923 ---------AACTGTCATTGAGCTCCCGGGAGCCCTATGGGAGGAGGGGCTCAGGGTCAC 2973
Db      2999 CTAAAGGAAAACTATCGTCGCGCTCCTGGGGGCCCTATGGGCAGAGGGGCTGCAGG---- 3054
Qy      2974 TCTTTTCAGGAAGCGCCAGCCCTAGGAACCCCAGAAAAGAGTGGTACCTAAGGCTGAAAT 3033
Db      3055 GTCTTTCAGGAAGCTCCAGCCCTAGGGACCCCAGAAAAGAGTGGCACCTATGGCTGAGTT 3114
Qy      3034 TGTTTTGCTGTTGCCAGGGGTGGGTATTTGAGAGTAAAACATTTTATTTCTTTTTAAAAA 3093
Db      3115 TGTTTTGCT------------GGTATTTGAGAGTAAAACATTTTATTTCTTTTTAAAAA 3161
Qy      3094 AAAAAAAAAAAAA 3106
Db      3162 AAAAAAAAAAAAA 3174
```

RESULT 5
AB037898

LOCUS	AB037898 2568 bp mRNA linear ROD 20-OCT-2001
DEFINITION	Rattus norvegicus MBSP mRNA for membrane bound serine protease, complete cds.
ACCESSION	AB037898
VERSION	AB037898.1 GI:9650963
KEYWORDS	membrane bound serine protease; MT-SP1.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1
AUTHORS	Satomi,S., Yamasaki,Y., Tsuzuki,S., Hitomi,Y., Iwanaga,T. and Fushiki,T.
TITLE	A role for membrane-type serine protease (MT-SP1) in intestinal epithelial turnover
JOURNAL	Biochem. Biophys. Res. Commun. 287 (4), 995-1002 (2001)
MEDLINE	21458307
PUBMED	11573963
REFERENCE	2 (bases 1 to 2568)
AUTHORS	Tsuzuki,S.
TITLE	Direct Submission
JOURNAL	Submitted (26-JAN-2000) Satoshi Tsuzuki, Kyoto University, Graduate School of Agriculture, Division of Food Science and Biotechnology,

Laboratory of Nutrition Chemistry; Oiwake-cho, Kitashirakawa,
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(E-mail:tkmone@kais.kais.kyoto-u.ac.jp, Tel:81-75-753-6263,
Fax:81-75-753-6264)

```
FEATURES             Location/Qualifiers
     source          1. .2568
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /tissue_type="jejunum"
                     /dev_stage="8 weeks old"
     gene            1. .2568
                     /gene="MBSP"
     CDS             1. .2568
                     /gene="MBSP"
                     /codon_start=1
                     /product="membrane bound serine protease"
                     /protein_id="BAB03502.1"
                     /db_xref="GI:9650964"
                     /translation="MGNNRGRKAGGGSQDFGAGLKYNSRLENMNGFEEGVEFLPVNNA
                     KQVEKRGPRRWVVMVAVVFSFLLLSLMAGLLVWHFHYRNVRIQKVFNGHLRITNENFL
                     DAYENSTSTEFISLASQVKEALKLMYSEVPVLGPYHKKSTVTAFSEGSVIAYYWSEFS
                     IPPHLEEEVDRAMAVERVVTLPPRARALKSFVLTSVVAFPIDPRMLQRTQDNSCSFAL
                     HARGRTVTRFTTPGFPNSPYPAHARCQWVLRGDADSVLSLTFRSFDVAPCDGHDSDLV
                     TVYDSLSPMEPHAVVRLCGTFSPSYNLTFLSSQNVFLVTLITNTDRRHPGFEATFFQL
                     PKMSSCGGLLSEAQGTFSSPYYPGHYPPNINCTWNIKVPNNRNVKVRFKLFYLVDPNI
                     PVGSCTKDYVEINGEKFCGERSQFVVSSNSSKITVHFHSDHSYTDTGFLAEYLSYDSN
                     DPCPGMFMCKTGRCIRKDLRCDGWADCPDYSDERHCRCNATHQFMCKNQFCKPLFWVC
                     DSVNDCGDGSDEEGCSCPAGSFKCSNGKCLPQSQQCNGKDDCGDGSDEASCDNVNAVS
                     CTKYTYRCQNGLCLNKGNPECDGKKDCSDGSDEKNCDCGLRSFTKQARVVGGTNADEG
                     EWPWQVSLHALGQGHLCGASLISPDWLVSAAHCFQDETIFKYSDHTMWTAFLGLLDQS
                     KRSASGVQEHKLKRIITHPSFNDFTFDYDIALLELEKPAEYSTVVRPICLPDNTHVFP
                     AGKAIWVTGWGHTKEGGTGALILQKGEIRVINQTTCEELLPQQITPRMMCVGFLSGGV
                     DSCQGDSGGPLSSVEKDGRIFQAGVVSWGEGCAQRNKPGVYTRIPEVRDWIKEQTGV"
ORIGIN
```

Query Match 74.5%; Score 2313.6; DB 10; Length 2568;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 2409; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

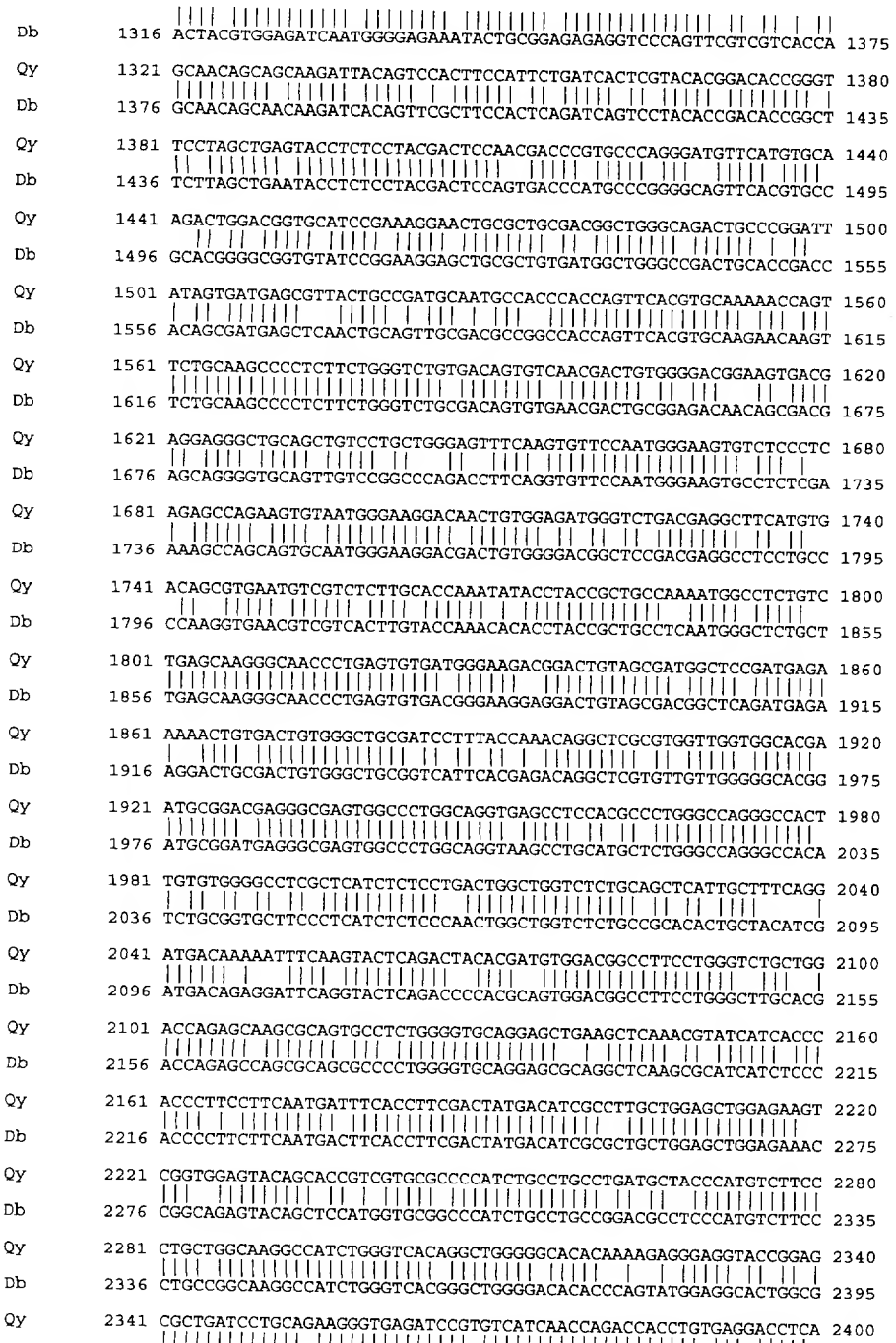
```
Qy      63 ATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTTCGGCGCGGGACTC 122
           ||||| | |||||||||||||||||||||||||||||||||||| |||||||||||||||
Db       1 ATGGGGAACAATCGGGGCCGCAAGGCCGGAGGGGGCTCGCAGGACTTCGGCGCCGGACTC 60

Qy     123 AAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGCCT 182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGCCT 120

Qy     183 GCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGCTGGTGGCA 242
           | |||||| |||||| | ||||||||||||||||||||||||||||||| |||||||
Db     121 GTGAACAACGCCAAGCAGGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGATGGTGGCC 180

Qy     243 GTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGTGTGGCACTTCCAT 302
           ||  ||||||||||| |||||||| |||||||| ||||||||||||||||||||||||||
Db     181 GTAGTGTTCAGCTTCCTTTTGCTCTCGCTCATGGCCGGCTTGCTGGTGTGGCACTTCCAT 240

Qy     303 TATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAAATGAGATC 362
           ||  ||||||||| || | ||||| |||||||||||||||||||||||| ||||||| |
Db     241 TACCGGAATGTGAGGATCCAAAAGGTCTTCAATGGCCATCTGAGGATCACGAATGAGAAC 300

Qy     363 TTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCTGGCCAGCCAGGTG 422
           |||||||| |||||||||||||||||||| ||||| |||||||||||||| |||||
Db     301 TTTCTGGACGCGTATGAGAACTCCACCTCCACGGAGTTCATCAGCCTGGCCAGTCAGGTT 360

Qy     423 AAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCCCTACCACAAGAAG 482
           |||||  |||||||||||| |||||| |||||||||||||||||||||||||||||||||
Db     361 AAGGAAGCGCTGAAGCTGATGTACAGTGAAGTCCCTGTCCTGGGCCCCTACCACAAGAAA 420

Qy     483 TCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTGGTCAGAGTTCAGC 542
           ||| |||||||||||||||||||||||| ||||| ||||||||||||||||||||||||
Db     421 TCGACTGTAACTGCCTTCAGTGAGGGCAGCGTCATTGCCTACTACTGGTCAGAGTTCAGC 480

Qy     543 ATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGAGCGAGTTGTAACA 602
           ||||||||||||||| |||||||||||||| |||||||||||||||||||||| |||
Db     481 ATCCCCCCACACCTGGAGGAAGAGGTTGATCGTGCCATGGCTGTGGAGCGAGTTGTCACA 540

Qy     603 TTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGTGGCCTTCCCC 662
           |||||||||||||||||||||||||||||||||||| ||||| || ||||||||||||||
Db     541 CTGCCACCCCGAGCACGGGCACTGAAATCCTTTGTGCTGACCTCTGTGGTGGCCTTCCCC 600

Qy     663 ATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTGCATGCC 722
           |||||||||||||||| ||||||||||||||||||||||||||||||||| ||||||
Db     601 ATTGACCCCAGAATGCTCCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTCCATGCC 660

Qy     723 CATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAGTCCCTACCCGGCG 782
           | |||      |||||||||| |||||||||||||||||||| |||||||||||||| ||
Db     661 CGTGGCAGGACAGTGACACGTTTCACTACCCCTGGCTTCCCTAACAGTCCCTACCCAGCT 720

Qy     783 CATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCTGTGCTGAGCCTCACCTTC 842
           |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db     721 CATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCGGACTCTGTGCTGAGCCTCACCTTC 780

Qy     843 CGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGTCACCGTGTATGAT 902
           |||||  |||||||||||  |||||||||  |||||  |||||||||||||||  ||||||||
Db     781 CGAAGTTTTGATGTCGCGCCCTGTGATGGGCATGACAGTGACCTGGTCACAGTGTATGAC 840

Qy     903 AGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCACCTTCTCACCCTCC 962
           |||||||||||||||| |||||  |||||||||||||||||||||||||||||| ||||||
Db     841 AGCCTGAGCCCCATGGAGCCCCATGCTGTGGTGCGGCTGTGTGGCACCTTCTCGCCCTCC 900

Qy     963 TACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCTGATAACCAATACT 1022
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     901 TACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTTACGCTGATAACCAACACT 960

Qy    1023 GACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCTGT 1082
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCTGT 1020

Qy    1083 GGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTATCCAGGCCACTAC 1142
           |||||  |||||||||  |||||||||  |||||||||||||||||||||||||||| |||
Db    1021 GGCGGCCTTTTGAGTGAAGCCCAAGGGACGTTTAGCAGCCCCTACTATCCAGGCCATTAC 1080

Qy    1143 CCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAGGTG 1202
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    1081 CCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAAGTA 1140

Qy    1203 CGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTCCTGCACCAAGGAC 1262
           |||||||||||||||||||||||||||||||| |||||||||  ||||||||||||||
Db    1141 CGCTTCAAACTCTTCTATCTGGTGGACCCCAACATACCAGTGGGAAGCTGCACCAAGGAC 1200

Qy    1263 TATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCAGC 1322
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    1201 TATGTGGAGATCAACGGGGAGAAGTTCTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCAGC 1260

Qy    1323 AACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCGTACACGGACACCGGGTTC 1382
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 AACAGCAGCAAGATTACGGTCCACTTCCATTCTGATCACTCGTACACGGACACCGGGTTC 1320

Qy    1383 CTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCAGGGATGTTCATGTGCAAG 1442
           || |||||||| |||||||||||||||| |||| |||||||||||||||||||| |||
Db    1321 CTGGCTGAGTATCTCTCCTACGACTCCAATGACCCATGCCCAGGGATGTTCATGTGTAAG 1380

Qy    1443 ACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGGGCAGACTGCCCGGATTAT 1502
           |||||||||||||||| |||||  |||||||||| |||||||| |||||||||| ||
Db    1381 ACTGGACGGTGCATCCGCAAGGACTTGCGCTGCGATGGCTGGGCTGACTGCCCGGACTAC 1440

Qy    1503 AGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTCACGTGCAAAAACCAGTTC 1562
           |||||||||| |||||||||| || |||||||||| |||||| |||||||||
Db    1441 AGTGATGAGCGTCACTGCCGATGCAACGCTACCCACCAGTTCATGTGCAAGAACCAGTTC 1500

Qy    1563 TGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGACGGAAGTGACGAG 1622
           |||||||||||||||||||||||||||||||||||||||||||||||||| || || ||| |
Db    1501 TGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGATGGCAGCGATGAG 1560
```

```
Qy       1623 GAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAATGGGAAGTGTCTCCCTCAG 1682
Db       1561 GAGGGCTGCAGTTGTCCTGCCGGGAGTTTCAAGTGTTCCAATGGGAAGTGTCTCCCTCAG 1620

Qy       1683 AGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCTGACGAGGCTTCATGTGAC 1742
Db       1621 AGCCAGCAGTGTAATGGGAAGGACGACTGTGGAGATGGGTCTGACGAGGCCTCATGTGAC 1680

Qy       1743 AGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAATGGCCTCTGTCTG 1802
Db       1681 AATGTGAATGCCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAATGGCCTCTGTCTG 1740

Qy       1803 AGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGCGATGGCTCCGATGAGAAA 1862
Db       1741 AACAAGGGCAACCCTGAGTGTGATGGGAAGAAGGACTGCAGCGATGGCTCGGATGAGAAA 1800

Qy       1863 AACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGCGTGGTTGGTGGCACGAAT 1922
Db       1801 AACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGAGTGGTTGGTGGCACGAAT 1860

Qy       1923 GCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCCCTGGGCCAGGGCCACTTG 1982
Db       1861 GCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCTCTGGGCCAGGGCCACTTG 1920

Qy       1983 TGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGGAT 2042
Db       1921 TGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCATTGCTTTCAGGAT 1980

Qy       2043 GACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCTGGGTCTGCTGGAC 2102
Db       1981 GAGACAATTTTCAAGTACTCAGACCACACCATGTGGACAGCCTTCCTCGGTCTGCTGGAC 2040

Qy       2103 CAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACGTATCATCACCCAC 2162
Db       2041 CAGAGCAAGCGCAGTGCCTCTGGGGTACAGGAGCACAAGCTCAAACGTATCATCACTCAC 2100

Qy       2163 CCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGAGCTGGAGAAGTCG 2222
Db       2101 CCTTCCTTTAATGACTTTACTTTCGACTATGACATTGCCTTGCTGGAGCTGGAGAAGCCG 2160

Qy       2223 GTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTACCCATGTCTTCCCT 2282
Db       2161 GCAGAGTACAGCACTGTCGTGCGCCCCATCTGCCTGCCTGATAATACCCATGTCTTTCCT 2220

Qy       2283 GCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGGAGGTACCGGAGCG 2342
Db       2221 GCCGGCAAAGCCATCTGGGTCACAGGCTGGGGCCACACGAAAGAAGGAGGAACTGGAGCA 2280

Qy       2343 CTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTGTGAGGACCTCATG 2402
Db       2281 CTGATCCTGCAGAAGGGAGAGATCCGGGTCATCAACCAGACCACCTGTGAGGAACTCTTG 2340

Qy       2403 CCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACTCC 2462
Db       2341 CCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGGGGGTGTGGACTCC 2400

Qy       2463 TGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAAGATGGGCGAATGTTCCAG 2522
Db       2401 TGCCAGGGTGATTCCGGTGGCCCCTTGTCCAGTGTGGAGAAAGATGGGCGAATCTTCCAG 2460

Qy       2523 GCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACACA 2582
Db       2461 GCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACACA 2520

Qy       2583 AGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATAG 2630
Db       2521 AGGATCCCTGAAGTTCGGGACTGGATCAAAGAGCAAACTGGGGTATAG 2568
```

RESULT 6
BC030532
LOCUS BC030532 3273 bp mRNA linear PRI 07-OCT-2003
DEFINITION Homo sapiens suppression of tumorigenicity 14 (colon carcinoma, matriptase, epithin), mRNA (cDNA clone MGC:40392 IMAGE:5213189), complete cds.
ACCESSION BC030532
VERSION BC030532.1 GI:20988874
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3273)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 3273)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 64 Row: k Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
FEATURES Location/Qualifiers
source 1. .3273
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40392 IMAGE:5213189"
/tissue_type="Blood, adult leukocytes"

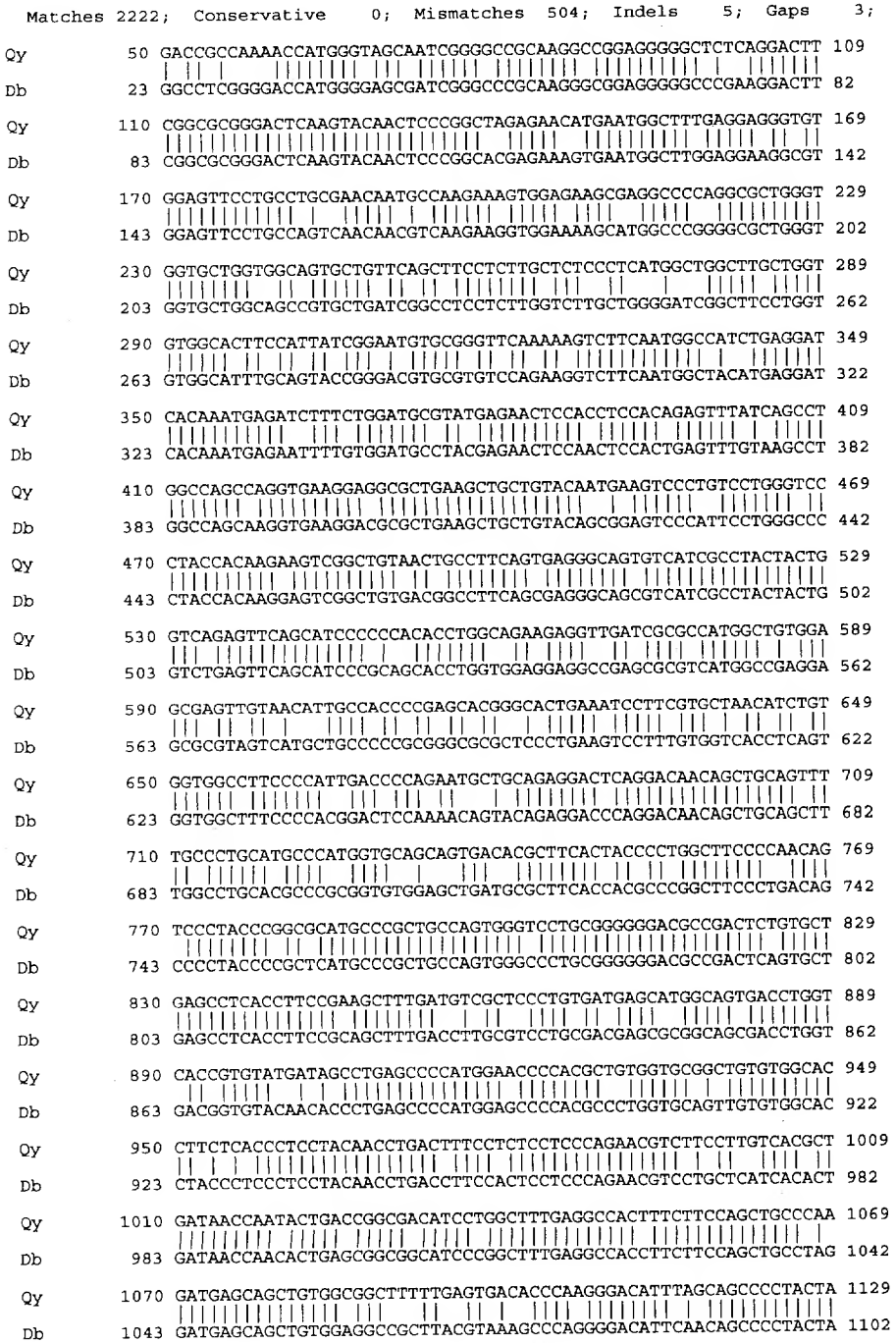
```
                     /clone_lib="NIH_MGC_118"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .3273
                     /gene="ST14"
                     /note="synonyms: SNC19, PRSS14, MTSP-1, TADG-15, MTSP1,
                     HAI, MT-SP1"
                     /db_xref="LocusID:6768"
                     /db_xref="MIM:606797"
     CDS             118. .2685
                     /codon_start=1
                     /product="matriptase"
                     /protein_id="AAH30532.1"
                     /db_xref="GI:20988875"
                     /db_xref="LocusID:6768"
                     /translation="MGSDRARKGGGGPKDFGAGLKYNSRHEKVNGLEEGVEFLPVNNV
                     KKVEKHGPGRWVVLAAVLIGLLLVLLGIGFLVWHLQYRDVRVQKVFNGYMRITNENFV
                     DAYENSNSTEFVSLASKVKDALKLLYSGVPFLGPYHKESAVTAFSEGSVIAYYWSEFS
                     IPQHLVEEAERVMAEERVVMLPPRARSLKSFVVTSVVAFPTDSKTVQRTQDNSCSFGL
                     HARGVELMRFTTPGFPDSPYPAHARCQWALRGDADSVLSLTFRSFDLASCDERGSDLV
                     TVYNTLSPMEPHALVQLCGTYPPSYNLTFHSSQNVLLITLITNTERRHPGFEATFFQL
                     PRMSSCGGRLRKAQGTFNSPYYPGHYPPNIDCTWNIEVPNNQHVKVRFKFFYLLEPGV
                     PAGTCPKDYVEINGEKYCGERSQFVVTSNSNKITVRFHSDQSYTDTGFLAEYLSYDSS
                     DPCPGQFTCRTGRCIRKELRCDGWADCTDHSDELNCSCDAGHQFTCKNKFCKPLFWVC
                     DSVNDCGDNSDEQGCSCPAQTFRCSNGKCLSKSQQCNGKDDCGDGSDEASCPKVNVVT
                     CTKHTYRCLNGLCLSKGNPECDGKEDCSDGSDEKDCDCGLRSFTRQARVVGGTDADEG
                     EWPWQVSLHALGQGHICGASLISPNWLVSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS
                     QRSAPGVQERRLKRIISHPFFNDFTFDYDIALLELEKPAEYSSMVRPICLPDASHVFP
                     AGKAIWVTGWGHTQYGGTGALILQKGEIRVINQTTCENLLPQQITPRMMCVGFLSGGV
                     DSCQGDSGGPLSSVEADGRIFQAGVVSWGDGCAQRNKPGVYTRLPLFRDWIKENTGV"
     misc_feature    796. .1110
                     /note="CUB; Region: CUB domain"
                     /db_xref="CDD:pfam00431"
     misc_feature    1135. .1449
                     /note="CUB; Region: CUB domain"
                     /db_xref="CDD:pfam00431"
     misc_feature    1477. .1581
                     /note="LDLa; Region: Low-density lipoprotein receptor
                     domain class A"
                     /db_xref="CDD:smart00192"
     misc_feature    1591. .1692
                     /note="LDLa; Region: Low-density lipoprotein receptor
                     domain class A"
                     /db_xref="CDD:smart00192"
     misc_feature    1690. .1800
                     /note="LDLa; Region: Low-density lipoprotein receptor
                     domain class A"
                     /db_xref="CDD:smart00192"
     misc_feature    1813. .1929
                     /note="LDLa; Region: Low-density lipoprotein receptor
                     domain class A"
                     /db_xref="CDD:smart00192"
     misc_feature    1957. .2664
                     /note="Tryp_SPc; Region: Trypsin-like serine protease"
                     /db_xref="CDD:smart00020"
ORIGIN

  Query Match             61.4%;  Score 1905.6;  DB 9;  Length 3273;
  Best Local Similarity   81.3%;  Pred. No. 0;
  Matches 2260;  Conservative    0;  Mismatches  514;  Indels    6;  Gaps    4;

Qy          1 CATGGTAGACGGCTGCCCGGAGGGACCACGCGTCTGAGACCGGCGATCGGACCGCCAAAA 60
              ||||| || |||| ||| | |||||| |||| | ||||||| |||| ||| || |    |
Db         57 CATGGGAGCCGGCCGCCGGCAGGGACGACGCCTGTGAGACCCGCGAGCGG-CCTCGGGGA 115

Qy         61 CCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTTCGGCGCGGGAC 120
              ||||||| ||| |||||| |||||||| |||||||||| | ||||||||||||||||||
Db        116 CCATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGGCCCGAAGGACTTCGGCGCGGGAC 175

Qy        121 TCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGTGGAGTTCCTGC 180
              |||||||||||||||||  |||||  |||||||||| || ||||||||||||||||||
Db        176 TCAAGTACAACTCCCGGCACGAGAAAGTGAATGGCTTGGAGGAAGGCGTGGAGTTCCTGC 235

Qy        181 CTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGTGGTGCTGGTGG 240
              | | ||||| | |||||| ||||| |||| |||||  ||||||||||||||||||| |
Db        236 CAGTCAACAACGTCAAGAAGGTGGAAAAGCATGGCCCGGGGCGCTGGGTGGTGCTGGCAG 295

Qy        241 CAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGTGTGGCACTTCC 300
              | |||||| || || |||||||| ||| ||  |  ||||| |||||||||| || |
Db        296 CCGTGCTGATCGGCCTCCTCTTGGTCTTGCTGGGGATCGGCTTCCTGGTGTGGCATTTGC 355

Qy        301 ATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAAATGAGA 360
              | || ||| | ||||| || || ||||||||||| |  |||||||||||||||||||
Db        356 AGTACCGGGACGTGCGTGTCCAGAAGGTCTTCAATGGCTACATGAGGATCACAAATGAGA 415

Qy        361 TCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCTGGCCAGCCAGG 420
              |||  ||||||| || |||||||||| |||||| |||||| | ||||||||||| |||
Db        416 ATTTTGTGGATGCCTACGAGAACTCCAACTCCACTGAGTTTGTAAGCCTGGCCAGCAAGG 475

Qy        421 TGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCCCTACCACAAGA 480
              ||||||| |||||||||||||||||||  | ||||||  |||||| ||||||||||||
Db        476 TGAAGGACGCGCTGAAGCTGCTGTACAGCGGAGTCCCATTCCTGGGCCCCTACCACAAGG 535

Qy        481 AGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTGGTCAGAGTTCA 540
              |||||||||| || |||||||| |||||||||||||||||||||||||||| |||||||
Db        536 AGTCGGCTGTGACGGCCTTCAGCGAGGGCAGCGTCATCGCCTACTACTGGTCTGAGTTCA 595

Qy        541 GCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGAGCGAGTTGTAA 600
              ||||||| |  |||||||  || ||||  || |||| |||||  | ||||||  || || |
Db        596 GCATCCCGCAGCACCTGGTGGAGGAGGCCGAGCGCGTCATGGCCGAGGAGCGCGTAGTCA 655

Qy        601 CATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGTGGTGGCCTTCC 660
              ||||  || || || ||  | ||||| ||||| ||| | || || |||||||| ||||
Db        656 TGCTGCCCCCGCGGGCGCGCTCCCTGAAGTCCTTTGTGGTCACCTCAGTGGTGGCTTTCC 715

Qy        661 CCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTTTGCCCTGCATG 720
              |||  |||  ||| ||    | |||||||| |||||||||||||||  |||| ||||||
Db        716 CCACGGACTCCAAAACAGTACAGAGGACCCAGGACAACAGCTGCAGCTTTGGCCTGCACG 775

Qy        721 CCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAGTCCCTACCCGG 780
              |||  ||||  |  |||  |||||||| || || ||||||||  |||| |||||||  |
Db        776 CCCGCGGTGTGGAGCTGATGCGCTTCACCACGCCCGGCTTCCCTGACAGCCCCTACCCCG 835

Qy        781 CGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGGACGCCGACTCTGTGCTGAGCCTCACCT 840
              | |||||||||||||||||| |||||||||||||||||||| ||||||||||||||||
Db        836 CTCATGCCCGCTGCCAGTGGGCCCTGCGGGGGGACGCCGACTCAGTGCTGAGCCTCACCT 895

Qy        841 TCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGTCACCGTGTATG 900
              |||| |||||||  | ||  |||| || ||||  |||||  |||||||| || |||||
Db        896 TCCGCAGCTTTGACCTTGCGTCCTGCGACGAGCGCGGCAGCGACCTGGTGACGGTGTACA 955

Qy        901 ATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCACCTTCTCACCCT 960
              | | |||||||||||||| ||||||||  |||||| | |||||||||||| | | ||||
Db        956 ACACCCTGAGCCCCATGGAGCCCCACGCCCTGGTGCAGTTGTGTGGCACCTACCCTCCCT 1015

Qy        961 CCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCTGATAACCAATA 1020
              |||||||||||||  |||| ||||||||||||||| | ||  |||| ||||||||||| |
Db       1016 CCTACAACCTGACCTTCCACTCCTCCCAGAACGTCCTGCTCATCACACTGATAACCAACA 1075

Qy       1021 CTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAAGATGAGCAGCT 1080
              |||| ||||| |||||  |||||||||||||||||||||||||||| | ||||||||||
Db       1076 CTGAGCGGCGGCATCCCGGCTTTGAGGCCACCTTCTTCCAGCTGCCTAGGATGAGCAGCT 1135

Qy       1081 GTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTATCCAGGCCACT 1140
              |||| |||  ||  ||  | |||| |||||||| | |||||||||||| ||||||||||
Db       1136 GTGGAGGCCGCTTACGTAAAGCCCAGGGGACATTCAACAGCCCCTACTACCCAGGCCACT 1195

Qy       1141 ACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCGGAACGTGAAGG 1200
              |||| |||||||| |||||||||||||| || |||||||||||||| | | |||||||
Db       1196 ACCCACCCAACATTGACTGCACATGGAACATTGAGGTGCCCAACAACCAGCATGTGAAGG 1255

Qy       1201 TGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTCCTGCACCAAGG 1260
              |||||||||| ||||||| ||| |||| ||| ||| || | |||| |||||  ||||||
Db       1256 TGCGCTTCAAATTCTTCTACCTGCTGGAGCCCGGCGTGCCTGCGGGCACCTGCCCCAAGG 1315

Qy       1261 ACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTTTGTGGTGAGCA 1320
```

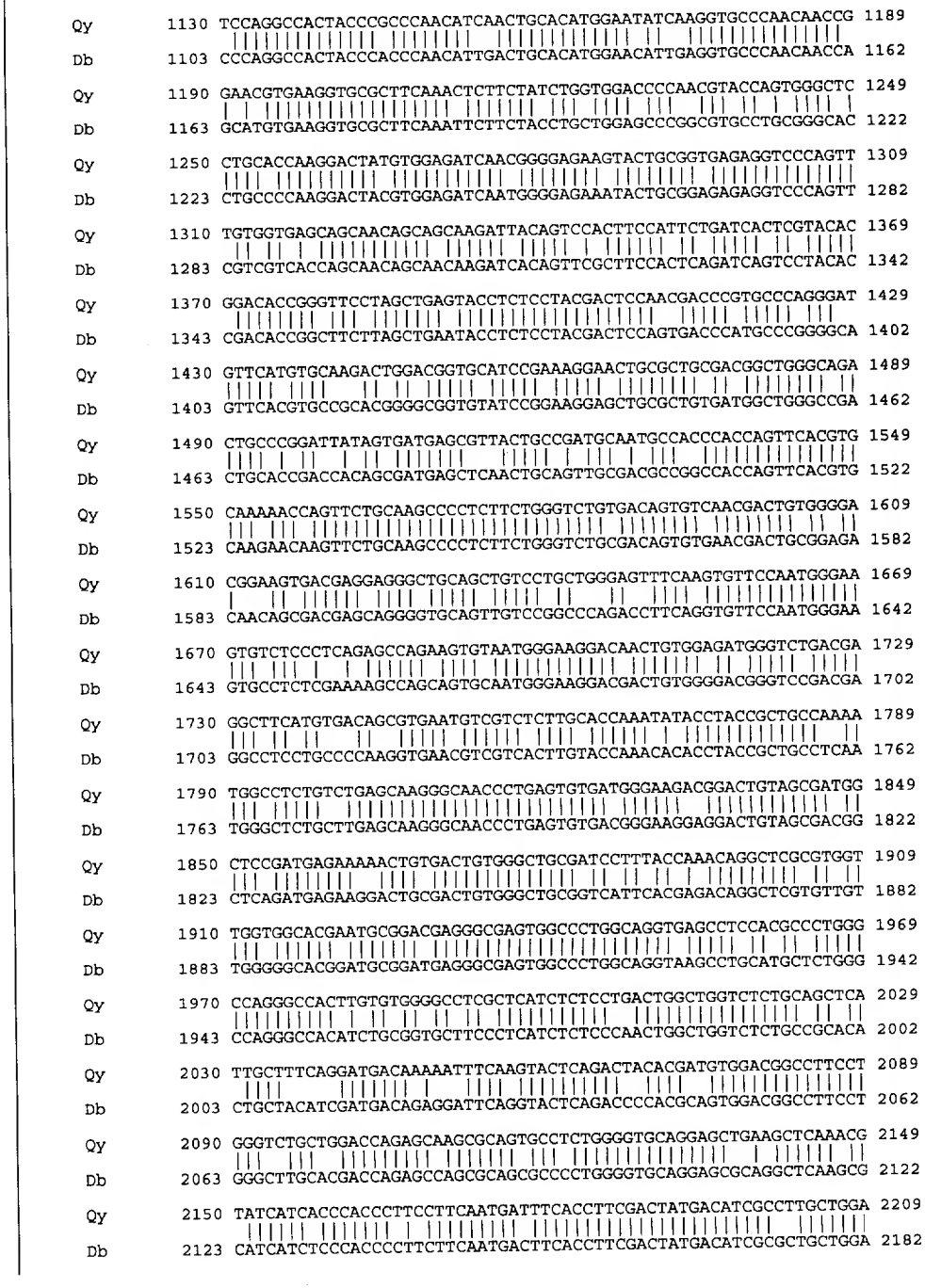


RESULT 7
BD274671
LOCUS BD274671 3149 bp DNA linear PAT 17-JUL-2003
DEFINITION Matriptase, serine protease and application of the same.
ACCESSION BD274671
VERSION BD274671.1 GI:33084439
KEYWORDS JP 2002539093-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3149)
AUTHORS Dickson,R.B., Lin,C., Johnson,M., Wang,S. and Enye,I.
TITLE Matriptase, serine protease and application of the same
JOURNAL Patent: JP 2002539093-A 2 19-NOV-2002;
GEORGETOWN UNIVERSITY
COMMENT OS Homo sapiens (human)
PN JP 2002539093-A/2
PD 19-NOV-2002
PF 10-MAR-2000 JP 2000603721
PR 12-MAR-1999 US 60/124006
PI ROBERT B DICKSON,CHEN-YONG LIN,MICHAEL JOHNSON,SHAOMENG WANG,
PI ISTVAN ENYEDY
PC A61K45/00,A61K35/78,A61K51/00,A61P1/00,A61P17/00,A61P35/00, PC
A61P43/00,
PC A61P43/00,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
,C12N9/64,C12N15/09,
PC C12Q1/37,G01N33/53,G01N33/577//C12P21/08,A61K49/02,C12N15/00,
PC C12N5/00
CC Matriptase, serine protease and application of the same FH
Key Location/Qualifiers
FT source 1. .3149
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1. .3149
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 60.8%; Score 1889.8; DB 6; Length 3149;
Best Local Similarity 81.4%; Pred. No. 0;

Matches 2222; Conservative 0; Mismatches 504; Indels 5; Gaps 3;





```
Qy       2210 GCTGGAGAAGTCGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTAC 2269
Db       2183 GCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGACGCCTC 2242

Qy       2270 CCATGTCTTCCCTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGG 2329
Db       2243 CCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGGGCTGGGGACACACCCAGTATGG 2302

Qy       2330 AGGTACCGGAGCGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTG 2389
Db       2303 AGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACCACCTG 2362

Qy       2390 TGAGGACCTCATGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGG 2449
Db       2363 CGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGATGTGCGTGGGCTTCCTCAGCGG 2422

Qy       2450 GGGTGTGGACTCCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAAGATGG 2509
Db       2423 CGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGTGGAGGCGGATGG 2482

Qy       2510 GCGAATGTTCCAGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCC 2569
Db       2483 GCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAACAAGCC 2542

Qy       2570 AGGCGTGTACACAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATA 2629
Db       2543 AGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAGAGAACACTGGGGTATA 2602

Qy       2630 GCAGCATGGACAGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGG 2689
Db       2603 GGGGCCGGG---GCCACCCAAATGTGTACACCTGCGGGGCCACCCATCGTCCACCCCAGT 2659

Qy       2690 ATACAGGAGAGGAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACAACCCA 2749
Db       2660 GTGCACG-CCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCC-CCAGAACATA 2717

Qy       2750 GACTGTGAACTGCATCCTTAGGACTCAGAGT 2780
Db       2718 CACTGTGAACTCAATCTCCAGGGCTCCAAAT 2748
```

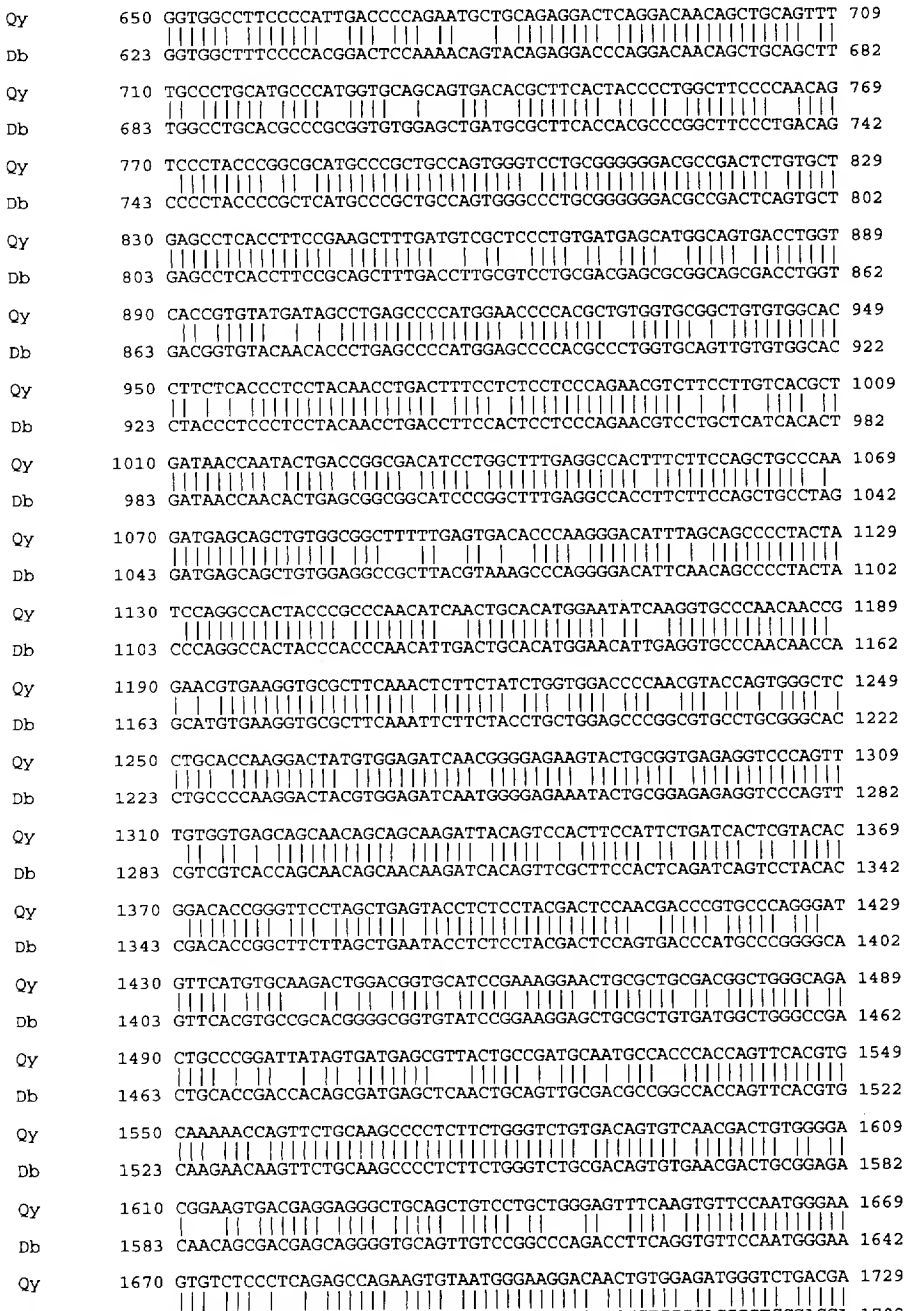
RESULT 8
AF118224
LOCUS AF118224 3149 bp mRNA linear PRI 30-DEC-1999
DEFINITION Homo sapiens matriptase mRNA, complete cds.
ACCESSION AF118224
VERSION AF118224.2 GI:6647301
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2955)
AUTHORS Lin,C.Y., Anders,J., Johnson,M., Sang,Q.A. and Dickson,R.B.
TITLE Molecular cloning of cDNA for matriptase, a matrix-degrading serine protease with trypsin-like activity
JOURNAL J. Biol. Chem. 274 (26), 18231-18236 (1999)
MEDLINE 99303581
PUBMED 10373424
REFERENCE 2 (bases 1 to 2955)
AUTHORS Lin,C.Y., Anders,J., Johnson,M., Sang,Q.A. and Dickson,R.B.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-1999) Lombardi Cancer Center, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington, DC 20007, USA
REFERENCE 3 (bases 1 to 3149)
AUTHORS Oberst,M.D., Dickson,R.B. and Lin,C.Y.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Lombardi Cancer Center, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington, DC 20007, USA
REMARK Sequence update by submitter
COMMENT On Dec 30, 1999 this sequence version replaced gi:5359674.

FEATURES Location/Qualifiers
source 1. .3149
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
CDS 36. .2603
/function="degrade extracellular matrix"
/note="serine protease with trypsin-like activity; contains C1r/s and LDL receptor domains"
/codon_start=1
/product="matriptase"
/protein_id="AAD42765.2"
/db_xref="GI:6647302"
/translation="MGSDRARKGGGGPKDFGAGLKYNSRHEKVNGLEEGVEFLPVNNV
KKVEKHGPGRWVVLAAVLIGLLLVLLGIGFLVWHLQYRDVRVQKVFNGYMRITNENFV
DAYENSNSTEFVSLASKVKDALKLLYSGVPFLGPYHKESAVTAFSEGSVIAYYWSEFS
IPQHLVEEAERVMAEERVVMLPPRARSLKSFVVTSVVAFPTDSKTVQRTQDNSCSFGL
HARGVELMRFTTPGFPDSPYPAHARCQWALRGDADSVLSLTFRSFDLASCDERGSDLV
TVYNTLSPMEPHALVQLCGTYPPSYNLTFHSSQNVLLITLITNTERRHPGFEATFFQL
PRMSSCGGRLRKAQGTFNSPYYPGHYPPNIDCTWNIEVPNNQHVKVRFKFFYLLEPGV
PAGTCPKDYVEINGEKYCGERSQFVVTSNSNKITVRFHSDQSYTDTGFLAEYLSYDSS
DPCPGQFTCRTGRCIRKELRCDGWADCTDHSDELNCSCDAGHQFTCKNKFCKPLFWVC
DSVNDCGDNSDEQGCSCPAQTFRCSNGKCLSKSQQCNGKDDCGDGSDEASCPKVNVVT
CTKHTYRCLNGLCLSKGNPECDGKEDCSDGSDEKDCDCGLRSFTRQARVVGGTDADEG
EWPWQVSLHALGQGHICGASLISPNWLVSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS
QRSAPGVQERRLKRIISHPFFNDFTFDYDIALLELEKPAEYSSMVRPICLPDASHVFP
AGKAIWVTGWGHTQYGGTGALILQKGEIRVINQTTCENLLPQQITPRMMCVGFLSGGV
DSCQGDSGGPLSSVEADGRIFQAGVVSWGDGCAQRNKPGVYTRLPLFRDWIKENTGV"

ORIGIN

Query Match 60.8%; Score 1889.8; DB 9; Length 3149;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 2222; Conservative 0; Mismatches 504; Indels 5; Gaps 3;

```
Qy         50 GACCGCCAAAACCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTT 109
Db         23 GGCCTCGGGGACCATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGGCCCGAAGGACTT 82

Qy        110 CGGCGCGGGACTCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGT 169
Db         83 CGGCGCGGGACTCAAGTACAACTCCCGGCACGAGAAAGTGAATGGCTTGGAGGAAGGCGT 142

Qy        170 GGAGTTCCTGCCTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGT 229
Db        143 GGAGTTCCTGCCAGTCAACAACGTCAAGAAGGTGGAAAAGCATGGCCCGGGGCGCTGGGT 202

Qy        230 GGTGCTGGTGGCAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGT 289
Db        203 GGTGCTGGCAGCCGTGCTGATCGGCCTCCTCTTGGTCTTGCTGGGGATCGGCTTCCTGGT 262

Qy        290 GTGGCACTTCCATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGAT 349
Db        263 GTGGCATTTGCAGTACCGGGACGTGCGTGTCCAGAAGGTCTTCAATGGCTACATGAGGAT 322

Qy        350 CACAAATGAGATCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCT 409
Db        323 CACAAATGAGAATTTTGTGGATGCCTACGAGAACTCCAACTCCACTGAGTTTGTAAGCCT 382

Qy        410 GGCCAGCCAGGTGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCC 469
Db        383 GGCCAGCAAGGTGAAGGACGCGCTGAAGCTGCTGTACAGCGGAGTCCCATTCCTGGGCCC 442

Qy        470 CTACCACAAGAAGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTG 529
Db        443 CTACCACAAGGAGTCGGCTGTGACGGCCTTCAGCGAGGGCAGCGTCATCGCCTACTACTG 502

Qy        530 GTCAGAGTTCAGCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGA 589
Db        503 GTCTGAGTTCAGCATCCCGCAGCACCTGGTGGAGGAGGCCGAGCGCGTCATGGCCGAGGA 562

Qy        590 GCGAGTTGTAACATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGT 649
Db        563 GCGCGTAGTCATGCTGCCCCCGCGGGCGCGCTCCCTGAAGTCCTTTGTGGTCACCTCAGT 622
```


Qy 650 GGTGGCCTTCCCCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTT 709
Db 623 GGTGGCTTTCCCCACGGACTCCAAAACAGTACAGAGGACCCAGGACAACAGCTGCAGCTT 682
Qy 710 TGCCCTGCATGCCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAG 769
Db 683 TGGCCTGCACGCCCGCGGTGTGGAGCTGATGCGCTTCACCACGCCCGGCTTCCCTGACAG 742
Qy 770 TCCCTACCCGGCGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCTGTGCT 829
Db 743 CCCCTACCCCGCTCATGCCCGCTGCCAGTGGGCCCTGCGGGGGGACGCCGACTCAGTGCT 802
Qy 830 GAGCCTCACCTTCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGT 889
Db 803 GAGCCTCACCTTCCGCAGCTTTGACCTTGCGTCCTGCGACGAGCGCGGCAGCGACCTGGT 862
Qy 890 CACCGTGTATGATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCAC 949
Db 863 GACGGTGTACAACACCCTGAGCCCCATGGAGCCCCACGCCCTGGTGCAGTTGTGTGGCAC 922
Qy 950 CTTCTCACCCTCCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCT 1009
Db 923 CTACCCTCCCTCCTACAACCTGACCTTCCACTCCTCCCAGAACGTCCTGCTCATCACACT 982
Qy 1010 GATAACCAATACTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAA 1069
Db 983 GATAACCAACACTGAGCGGCGGCATCCCGGCTTTGAGGCCACCTTCTTCCAGCTGCCTAG 1042
Qy 1070 GATGAGCAGCTGTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTA 1129
Db 1043 GATGAGCAGCTGTGGAGGCCGCTTACGTAAAGCCCAGGGGACATTCAACAGCCCCTACTA 1102
Qy 1130 TCCAGGCCACTACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCG 1189
Db 1103 CCCAGGCCACTACCCACCCAACATTGACTGCACATGGAACATTGAGGTGCCCAACAACCA 1162
Qy 1190 GAACGTGAAGGTGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTC 1249
Db 1163 GCATGTGAAGGTGCGCTTCAAATTCTTCTACCTGCTGGAGCCCGGCGTGCCTGCGGGCAC 1222
Qy 1250 CTGCACCAAGGACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTT 1309
Db 1223 CTGCCCCAAGGACTACGTGGAGATCAATGGGGAGAAATACTGCGGAGAGAGGTCCCAGTT 1282
Qy 1310 TGTGGTGAGCAGCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCGTACAC 1369
Db 1283 CGTCGTCACCAGCAACAGCAACAAGATCACAGTTCGCTTCCACTCAGATCAGTCCTACAC 1342
Qy 1370 GGACACCGGGTTCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCAGGGAT 1429
Db 1343 CGACACCGGCTTCTTAGCTGAATACCTCTCCTACGACTCCAGTGACCCATGCCCGGGGCA 1402
Qy 1430 GTTCATGTGCAAGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGGGCAGA 1489
Db 1403 GTTCACGTGCCGCACGGGGCGGTGTATCCGGAAGGAGCTGCGCTGTGATGGCTGGGCCGA 1462
Qy 1490 CTGCCCGGATTATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTCACGTG 1549
Db 1463 CTGCACCGACCACAGCGATGAGCTCAACTGCAGTTGCGACGCCGGCCACCAGTTCACGTG 1522
Qy 1550 CAAAAACCAGTTCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGA 1609
Db 1523 CAAGAACAAGTTCTGCAAGCCCCTCTTCTGGGTCTGCGACAGTGTGAACGACTGCGGAGA 1582
Qy 1610 CGGAAGTGACGAGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAATGGGAA 1669
Db 1583 CAACAGCGACGAGCAGGGGTGCAGTTGTCCGGCCCAGACCTTCAGGTGTTCCAATGGGAA 1642
Qy 1670 GTGTCTCCCTCAGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCTGACGA 1729
Db 1643 GTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTGGGGACGGGTCCGACGA 1702
Qy 1730 GGCTTCATGTGACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAA 1789
Db 1703 GGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCTACCGCTGCCTCAA 1762
Qy 1790 TGGCCTCTGTCTGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGCGATGG 1849
Db 1763 TGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGGAGGACTGTAGCGACGG 1822
Qy 1850 CTCCGATGAGAAAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGCGTGGT 1909
Db 1823 CTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGTCATTCACGAGACAGGCTCGTGTTGT 1882
Qy 1910 TGGTGGCACGAATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCCCTGGG 1969
Db 1883 TGGGGGCACGGATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAAGCCTGCATGCTCTGGG 1942
Qy 1970 CCAGGGCCACTTGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCA 2029
Db 1943 CCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGCTGGTCTCTGCCGCACA 2002
Qy 2030 TTGCTTTCAGGATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCT 2089
Db 2003 CTGCTACATCGATGACAGAGGATTCAGGTACTCAGACCCCACGCAGTGGACGGCCTTCCT 2062
Qy 2090 GGGTCTGCTGGACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACG 2149
Db 2063 GGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGGAGCGCAGGCTCAAGCG 2122
Qy 2150 TATCATCACCCACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGA 2209
Db 2123 CATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATGACATCGCGCTGCTGGA 2182
Qy 2210 GCTGGAGAAGTCGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTAC 2269
Db 2183 GCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGACGCCTC 2242
Qy 2270 CCATGTCTTCCCTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAAGAGGG 2329
Db 2243 CCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGGGCTGGGGACACACCCAGTATGG 2302
Qy 2330 AGGTACCGGAGCGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACCACCTG 2389
Db 2303 AGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACCACCTG 2362
Qy 2390 TGAGGACCTCATGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTCAGTGG 2449
Db 2363 CGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGATGTGCGTGGGCTTCCTCAGCGG 2422
Qy 2450 GGGTGTGGACTCCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAAGATGG 2509
Db 2423 CGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGTGGAGGCGGATGG 2482
Qy 2510 GCGAATGTTCCAGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAACAAGCC 2569
Db 2483 GCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAACAAGCC 2542
Qy 2570 AGGCGTGTACACAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGGGTATA 2629
Db 2543 AGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAGAGAACACTGGGGTATA 2602
Qy 2630 GCAGCATGGACAGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACACCTGG 2689
Db 2603 GGGGCCGGG---GCCACCCAAATGTGTACACCTGCGGGGCCACCCATCGTCCACCCCAGT 2659
Qy 2690 ATACAGGAGAGGAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACAACCCA 2749
Db 2660 GTGCACG-CCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCC-CCAGAACATA 2717
Qy 2750 GACTGTGAACTGCATCCTTAGGACTCAGAGT 2780
Db 2718 CACTGTGAACTCAATCTCCAGGGCTCCAAAT 2748

RESULT 9
AF133086
LOCUS AF133086 3120 bp mRNA linear PRI 01-OCT-1999
DEFINITION Homo sapiens membrane-type serine protease 1 mRNA, complete cds.
ACCESSION AF133086
VERSION AF133086.1 GI:6002713
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3120)
AUTHORS Takeuchi,T., Shuman,M.A. and Craik,C.S.
TITLE Reverse biochemistry: use of macromolecular protease inhibitors to dissect complex biological processes and identify a membrane-type serine protease in epithelial cancer and normal tissue
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11054-11061 (1999)
MEDLINE 99432178
PUBMED 10500122
REFERENCE 2 (bases 1 to 3120)
AUTHORS Takeuchi,T., Shuman,M.A. and Craik,C.S.
TITLE Direct Submission
JOURNAL Submitted (04-MAR-1999) Dept. Pharm. Chem., University of California, San Francisco, 513 Parnassus Ave., Box 0446, San Francisco, CA 94143-0446, USA
FEATURES Location/Qualifiers
source 1. .3120
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="PC-3"
/cell_type="prostate carcinoma"
CDS 37. .2604
/note="MT-SP1; chymotrypsin clan serine protease"
/codon_start=1
/product="membrane-type serine protease 1"
/protein_id="AAF00109.1"
/db_xref="GI:6002714"
/translation="MGSDRARKGGGGPKDFGAGLKYNSRHEKVNGLEEGVEFLPVNNV
KKVEKHGPGRWVVLAAVLIGLLLVLLGIGFLVWHLQYRDVRVQKVFNGYMRITNENFV
DAYENSNSTEFVSLASKVKDALKLLYSGVPFLGPYHKESAVTAFSEGSVIAYYWSEFS
IPQHLVEEAERVMAEERVVMLPPRARSLKSFVVTSVVAFPTDSKTVQRTQDNSCSFGL
HARGVELMRFTTPGFPDSPYPAHARCQWALRGDADSVLSLTFRSFDLASCDERGSDLV
TVYNTLSPMEPHALVQLCGTYPPSYNLTFHSSQNVLLITLITNTERRHPGFEATFFQL
PRMSSCGGRLRKAQGTFNSPYYPGHYPPNIDCTWNIEVPNNQHVKVRFKFFYLLEPGV
PAGTCPKDYVEINGEKYCGERSQFVVTSNSNKITVRFHSDQSYTDTGFLAEYLSYDSS
DPCPGQFTCRTGRCIRKELRCDGWADCTDHSDELNCSCDAGHQFTCKNKFCKPLFWVC
DSVNDCGDNSDEQGCSCPAQTFRCSNGKCLSKSQQCNGKDDCGDGSDEASCPKVNVVT
CTKHTYRCLNGLCLSKGNPECDGKEDCSDGSDEKDCDCGLRSFTRQARVVGGTDADEG
EWPWQVSLHALGQGHICGASLISPNWLVSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS
QRSAPGVQERRLKRIISHPFFNDFTFDYDIALLELEKPAEYSSMVRPICLPDASHVFP
AGKAIWVTGWGHTQYGGTGALILQKGEIRVINQTTCENLLPQQITPRMMCVGFLSGGV
DSCQGDSGGPLSSVEADGRIFQAGVVSWGDGCAQRNKPGVYTRLPLFRDWIKENTGV"
ORIGIN

Query Match 60.7%; Score 1884.8; DB 9; Length 3120;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 447; Indels 1; Gaps 1;

```
Qy        34 CTGAGACCGGCGATCGGACCGCCAAAACCATGGGTAGCAATCGGGGCCGCAAGGCCGGAG 93
             ||||||| |||| ||| || |    |||||||| ||| |||||| |||||||| |||||
Db         9 CTGAGACCCGCGAGCGG-CCTCGGGGACCATGGGGAGCGATCGGGCCCGCAAGGGCGGAG 67

Qy        94 GGGGCTCTCAGGACTTCGGCGCGGGACTCAAGTACAACTCCCGGCTAGAGAACATGAATG 153
             ||||| | ||||||||||||||||||||||||||||||||||||||| ||||| ||||||
Db        68 GGGGCCCGAAGGACTTCGGCGCGGGACTCAAGTACAACTCCCGGCACGAGAAAGTGAATG 127

Qy       154 GCTTTGAGGAGGGTGTGGAGTTCCTGCCTGCGAACAATGCCAAGAAAGTGGAGAAGCGAG 213
             |||| ||||| || |||||||||||||| | ||||| | |||||| ||||| |||| |
Db       128 GCTTGGAGGAAGGCGTGGAGTTCCTGCCAGTCAACAACGTCAAGAAGGTGGAAAAGCATG 187

Qy       214 GCCCCAGGCGCTGGGTGGTGCTGGTGGCAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCA 273
             ||||  |||||||||||||||||  || |||||| || || |||||||| |||  ||
```

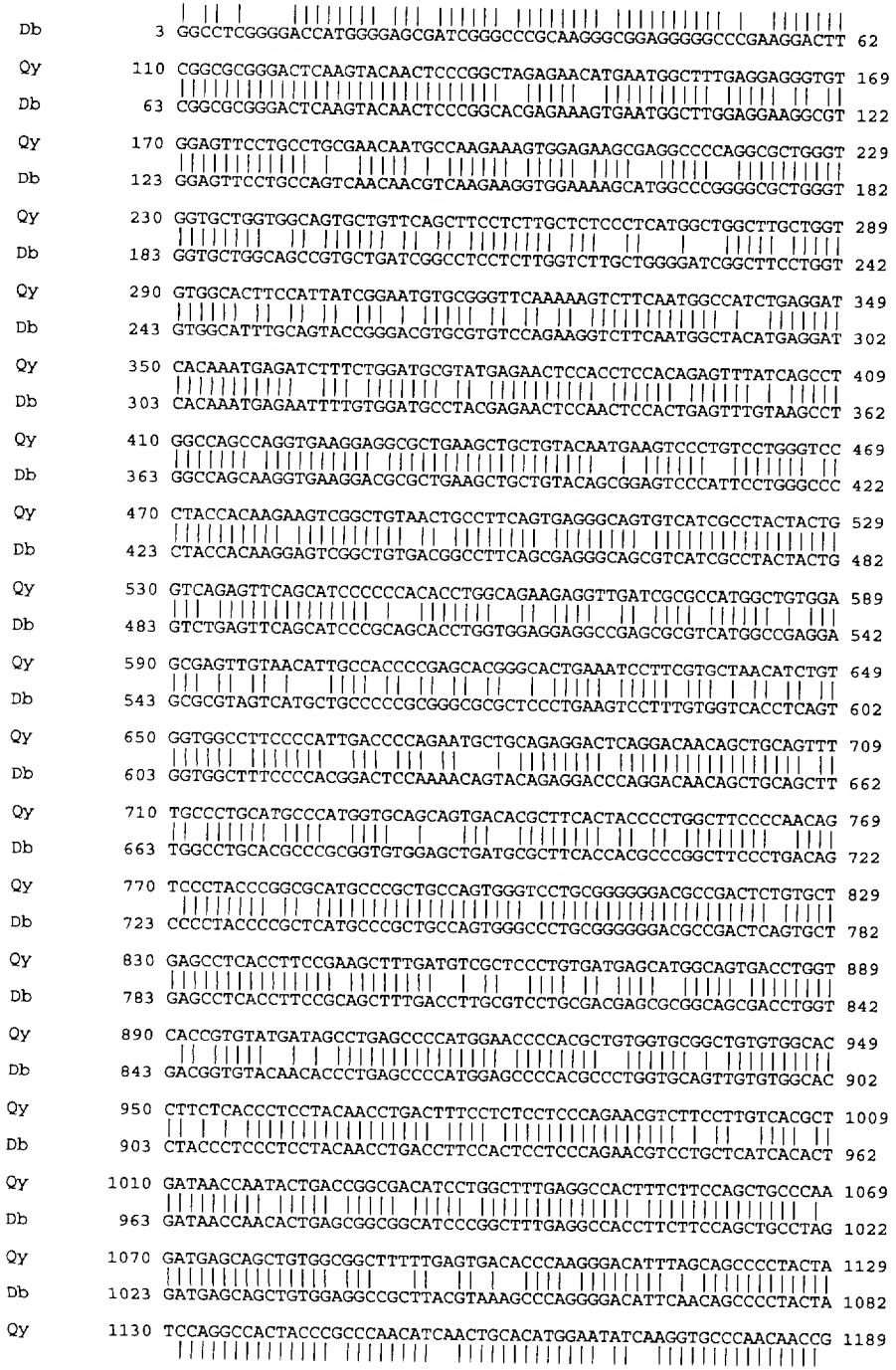
```
Db       188 GCCCGGGGCGCTGGGTGGTGCTGGCAGCCGTGCTGATCGGCCTCCTCTTGGTCTTGCTGG 247

Qy       274 TGGCTGGCTTGCTGGTGTGGCACTTCCATTATCGGAATGTGCGGGTTCAAAAAGTCTTCA 333
             |    ||||| |||||||||||| || || || ||| | ||||| || || || |||||||
Db       248 GGATCGGCTTCCTGGTGTGGCATTTGCAGTACCGGGACGTGCGTGTCCAGAAGGTCTTCA 307

Qy       334 ATGGCCATCTGAGGATCACAAATGAGATCTTTCTGGATGCGTATGAGAACTCCACCTCCA 393
             ||||| | |||||||||||||||||||  ||| ||||||| || |||||||||| |||||
Db       308 ATGGCTACATGAGGATCACAAATGAGAATTTTGTGGATGCCTACGAGAACTCCAACTCCA 367

Qy       394 CAGAGTTTATCAGCCTGGCCAGCCAGGTGAAGGAGGCGCTGAAGCTGCTGTACAATGAAG 453
             | |||||| | |||||||||||| |||||||||| |||||||||||||||||||| | ||
Db       368 CTGAGTTTGTAAGCCTGGCCAGCAAGGTGAAGGACGCGCTGAAGCTGCTGTACAGCGGAG 427

Qy       454 TCCCTGTCCTGGGTCCCTACCACAAGAAGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTG 513
             ||||  ||||||| |||||||||||| |||||||||| || |||||||| ||||||| |
Db       428 TCCCATTCCTGGGCCCCTACCACAAGGAGTCGGCTGTGACGGCCTTCAGCGAGGGCAGCG 487

Qy       514 TCATCGCCTACTACTGGTCAGAGTTCAGCATCCCCCCACACCTGGCAGAAGAGGTTGATC 573
             |||||||||||||||||||| |||||||||||| | |||||||  || ||||  || | |
Db       488 TCATCGCCTACTACTGGTCTGAGTTCAGCATCCCGCAGCACCTGGTGGAGGAGGCCGAGC 547

Qy       574 GCGCCATGGCTGTGGAGCGAGTTGTAACATTGCCACCCCGAGCACGGGCACTGAAATCCT 633
             ||| |||||| | |||||| || || |  |||| || || || ||  | ||||| ||||
Db       548 GCGTCATGGCCGAGGAGCGCGTAGTCATGCTGCCCCCGCGGGCGCGCTCCCTGAAGTCCT 607

Qy       634 TCGTGCTAACATCTGTGGTGGCCTTCCCCATTGACCCCAGAATGCTGCAGAGGACTCAGG 693
             | ||| | || || |||||||| ||||||| ||| ||| ||   | |||||||| ||||
Db       608 TTGTGGTCACCTCAGTGGTGGCTTTCCCCACGGACTCCAAAACAGTACAGAGGACCCAGG 667

Qy       694 ACAACAGCTGCAGTTTTGCCCTGCATGCCCATGGTGCAGCAGTGACACGCTTCACTACCC 753
             |||||||||||||| |||| |||||| |||| |||| |   |||  |||||||| || |
Db       668 ACAACAGCTGCAGCTTTGGCCTGCACGCCCGCGGTGTGGAGCTGATGCGCTTCACCACGC 727

Qy       754 CTGGCTTCCCCAACAGTCCCTACCCGGCGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGG 813
             | |||||||| |||| |||||||| || |||||||||||||||||| ||||||||||
Db       728 CCGGCTTCCCTGACAGCCCCTACCCCGCTCATGCCCGCTGCCAGTGGGCCCTGCGGGGGG 787

Qy       814 ACGCCGACTCTGTGCTGAGCCTCACCTTCCGAAGCTTTGATGTCGCTCCCTGTGATGAGC 873
             ||||||||| |||||||||||||||||||| |||||||| | || |||| || ||||
Db       788 ACGCCGACTCAGTGCTGAGCCTCACCTTCCGCAGCTTTGACCTTGCGTCCTGCGACGAGC 847

Qy       874 ATGGCAGTGACCTGGTCACCGTGTATGATAGCCTGAGCCCCATGGAACCCCACGCTGTGG 933
             |||| |||||||| || ||||| | | |||||||||||||||| ||||||| |||
Db       848 GCGGCAGCGACCTGGTGACGGTGTACAACACCCTGAGCCCCATGGAGCCCCACGCCCTGG 907

Qy       934 TGCGGCTGTGTGGCACCTTCTCACCCTCCTACAACCTGACTTTCCTCTCCTCCCAGAACG 993
             ||| | |||||||||||| | | |||||||||||||||||| |||| ||||||||||||||
Db       908 TGCAGTTGTGTGGCACCTACCCTCCCTCCTACAACCTGACCTTCCACTCCTCCCAGAACG 967

Qy       994 TCTTCCTTGTCACGCTGATAACCAATACTGACCGGCGACATCCTGGCTTTGAGGCCACTT 1053
             || | ||  |||| |||||||||| ||||| ||||| ||||| |||||||||||||| |
Db       968 TCCTGCTCATCACACTGATAACCAACACTGAGCGGCGGCATCCCGGCTTTGAGGCCACCT 1027

Qy      1054 TCTTCCAGCTGCCCAAGATGAGCAGCTGTGGCGGCTTTTTGAGTGACACCCAAGGGACAT 1113
             |||||||||||||| | |||||||||||||| |||   ||  || | |||| |||||||
Db      1028 TCTTCCAGCTGCCTAGGATGAGCAGCTGTGGAGGCCGCTTACGTAAAGCCCAGGGGACAT 1087

Qy      1114 TTAGCAGCCCCTACTATCCAGGCCACTACCCGCCCAACATCAACTGCACATGGAATATCA 1173
             | | |||||||||||| |||||||||||||| |||||||  |||||||||||||| ||
Db      1088 TCAACAGCCCCTACTACCCAGGCCACTACCCACCCAACATTGACTGCACATGGAACATTG 1147

Qy      1174 AGGTGCCCAACAACCGGAACGTGAAGGTGCGCTTCAAACTCTTCTATCTGGTGGACCCCA 1233
             |||||||||||||| | |||||||||||||||||| ||||||| ||| |||| |||
Db      1148 AGGTGCCCAACAACCAGCATGTGAAGGTGCGCTTCAAATTCTTCTACCTGCTGGAGCCCG 1207

Qy      1234 ACGTACCAGTGGGCTCCTGCACCAAGGACTATGTGGAGATCAACGGGGAGAAGTACTGCG 1293
             ||| || | |||| ||||| |||||||||||| |||||||| |||||||| |||||||
Db      1208 GCGTGCCTGCGGGCACCTGCCCCAAGGACTACGTGGAGATCAATGGGGAGAAATACTGCG 1267

Qy      1294 GTGAGAGGTCCCAGTTTGTGGTGAGCAGCAACAGCAGCAAGATTACAGTCCACTTCCATT 1353
             | |||||||||||||| || || | |||||||||| |||||| ||||| | |||||| |
Db      1268 GAGAGAGGTCCCAGTTCGTCGTCACCAGCAACAGCAACAAGATCACAGTTCGCTTCCACT 1327
```

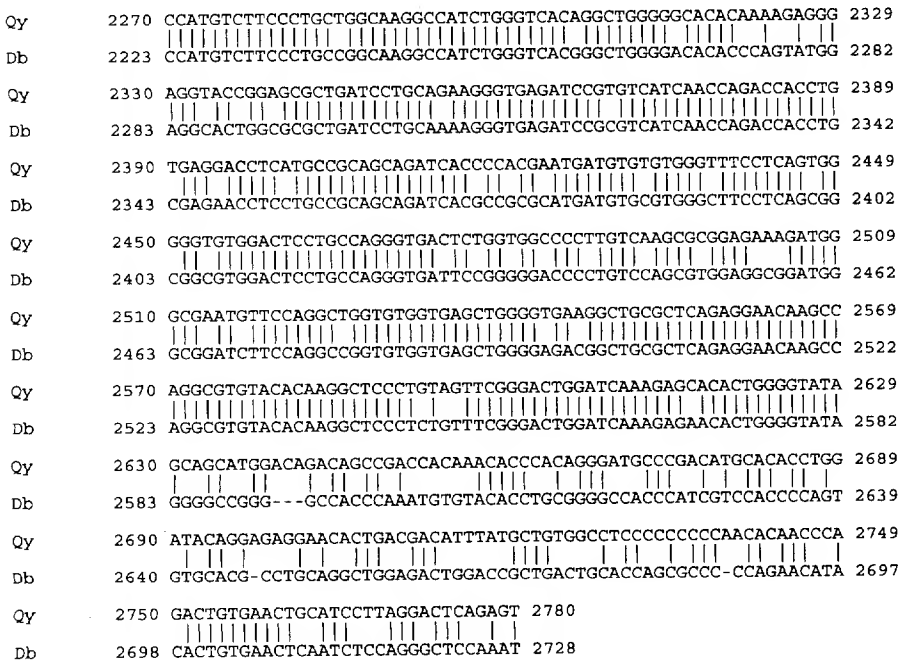
```
Qy       1354 CTGATCACTCGTACACGGACACCGGGTTCCTAGCTGAGTACCTCTCCTACGACTCCAACG 1413
Db       1328 CAGATCAGTCCTACACCGACACCGGCTTCTTAGCTGAATACCTCTCCTACGACTCCAGTG 1387

Qy       1414 ACCCGTGCCCAGGGATGTTCATGTGCAAGACTGGACGGTGCATCCGAAAGGAACTGCGCT 1473
Db       1388 ACCCATGCCCGGGGCAGTTCACGTGCCGCACGGGGCGGTGTATCCGGAAGGAGCTGCGCT 1447

Qy       1474 GCGACGGCTGGGCAGACTGCCCGGATTATAGTGATGAGCGTTACTGCCGATGCAATGCCA 1533
Db       1448 GTGATGGCTGGGCCGACTGCACCGACCACAGCGATGAGCTCAACTGCAGTTGCGACGCCG 1507

Qy       1534 CCCACCAGTTCACGTGCAAAAACCAGTTCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTG 1593
Db       1508 GCCACCAGTTCACGTGCAAGAACAAGTTCTGCAAGCCCCTCTTCTGGGTCTGCGACAGTG 1567

Qy       1594 TCAACGACTGTGGGGACGGAAGTGACGAGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCA 1653
Db       1568 TGAACGACTGCGGAGACAACAGCGACGAGCAGGGGTGCAGTTGTCCGGCCCAGACCTTCA 1627

Qy       1654 AGTGTTCCAATGGGAAGTGTCTCCCTCAGAGCCAGAAGTGTAATGGGAAGGACAACTGTG 1713
Db       1628 GGTGTTCCAATGGGAAGTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTG 1687

Qy       1714 GAGATGGGTCTGACGAGGCTTCATGTGACAGCGTGAATGTCGTCTCTTGCACCAAATATA 1773
Db       1688 GGGACGGGTCCGACGAGGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACA 1747

Qy       1774 CCTACCGCTGCCAAAATGGCCTCTGTCTGAGCAAGGGCAACCCTGAGTGTGATGGGAAGA 1833
Db       1748 CCTACCGCTGCCTCAATGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGG 1807

Qy       1834 CGGACTGTAGCGATGGCTCCGATGAGAAAAACTGTGACTGTGGGCTGCGATCCTTTACCA 1893
Db       1808 AGGACTGTAGCGACGGCTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGTCATTCACGA 1867

Qy       1894 AACAGGCTCGCGTGGTTGGTGGCACGAATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGA 1953
Db       1868 GACAGGCTCGTGTTGTTGGGGGCACGGATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAA 1927

Qy       1954 GCCTCCACGCCCTGGGCCAGGGCCACTTGTGTGGGGCCTCGCTCATCTCTCCTGACTGGC 2013
Db       1928 GCCTGCATGCTCTGGGCCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGC 1987

Qy       2014 TGGTCTCTGCAGCTCATTGCTTTCAGGATGACAAAAATTTCAAGTACTCAGACTACACGA 2073
Db       1988 TGGTCTCTGCCGCACACTGCTACATCGATGACAGAGGATTCAGGTACTCAGACCCCACGC 2047

Qy       2074 TGTGGACGGCCTTCCTGGGTCTGCTGGACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGG 2133
Db       2048 AGTGGACGGCCTTCCTGGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGG 2107

Qy       2134 AGCTGAAGCTCAAACGTATCATCACCCACCCTTCCTTCAATGATTTCACCTTCGACTATG 2193
Db       2108 AGCGCAGGCTCAAGCGCATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATG 2167

Qy       2194 ACATCGCCTTGCTGGAGCTGGAGAAGTCGGTGGAGTACAGCACCGTCGTGCGCCCCATCT 2253
Db       2168 ACATCGCGCTGCTGGAGCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCT 2227

Qy       2254 GCCTGCCTGATGCTACCCATGTCTTCCCTGCTGGCAAGGCCATCTGGGTCACAGGCTGGG 2313
Db       2228 GCCTGCCGGACGCCTCCCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGGGCTGGG 2287

Qy       2314 GGCACACAAAAGAGGGAGGTACCGGAGCGCTGATCCTGCAGAAGGGTGAGATCCGTGTCA 2373
Db       2288 GACACACCCAGTATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCA 2347

Qy       2374 TCAACCAGACCACCTGTGAGGACCTCATGCCGCAGCAGATCACCCCACGAATGATGTGTG 2433
Db       2348 TCAACCAGACCACCTGCGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGATGTGCG 2407

Qy       2434 TGGGTTTCCTCAGTGGGGGTGTGGACTCCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAA 2493
Db       2408 TGGGCTTCCTCAGCGGCGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCA 2467

Qy       2494 GCGCGGAGAAAGATGGGCGAATGTTCCAGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCG 2553
Db       2468 GCGTGGAGGCGGATGGGCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCG 2527

Qy       2554 CTCAGAGGAACAAGCCAGGCGTGTACACAAGGCTCCCTGTAGTTCGGGACTGGATCAAAG 2613
Db       2528 CTCAGAGGAACAAGCCAGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAG 2587

Qy       2614 AGCACACTGGGGTATAGCAGCATGGACAGACA 2645
Db       2588 AGAACACTGGGGTATAGGGGCCGGGGCACCCA 2619
```

RESULT 10
AB030036

```
LOCUS       AB030036               3128 bp    mRNA    linear   PRI 06-JAN-2001
DEFINITION  Homo sapiens mRNA for prostamin, complete cds.
ACCESSION   AB030036
VERSION     AB030036.1  GI:12249014
KEYWORDS    prostamin.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Yamaguchi,N. and Mitsui,S.
  TITLE     Molecular cloning of a novel transmembrane serine protease
            expressed in human prostate
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3128)
  AUTHORS   Yamaguchi,N. and Mitsui,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUL-1999) Nozomi Yamaguchi, Kyoto Prefectural
            University of Medicine, Res. Ins. Geriatrics; Kawaramachi Hirokoji,
            Kyoto, Kyoto 602-8566, Japan (E-mail:nozomi@koto.kpu-m.ac.jp,
            Tel:81-75-251-5848, Fax:81-75-251-5848)
FEATURES             Location/Qualifiers
     source          1. .3128
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="prostate"
     CDS             16. .2583
                     /codon_start=1
                     /product="prostamin"
                     /protein_id="BAB20376.1"
                     /db_xref="GI:12249015"
                     /translation="MGSDRARKGGGGPKDFGAGLKYNSRHEKVNGLEEGVEFLPVNNV
                     KKVEKHGPGRWVVLAAVLIGLLLVLLGIGFLVWHLQYRDVRVQKVFNGYMRITNENFV
                     DAYENSNSTEFVSLASKVKDALKLLYSGVPFLGPYHKESAVTAFSEGSVIAYYWSEFS
                     IPQHLVEEAERVMAEERVVMLPPRARSLKSFVVTSVVAFPTDSKTVQRTQDNSCSFGL
                     HARGVELMRFTTPGFPDSPYPAHARCQWALRGDADSVLSLTFRSFDLASCDERGSDLV
                     TVYNTLSPMEPHALVQLCGTYPPSYNLTFHSSQNVLLITLITNTERRHPGFEATFFQL
                     PRMSSCGGRLRKAQGTFNSPYYPGHYPPNIDCTWNIEVPNNQHVKVRFKFFYLLEPGV
                     PAGTCPKDYVEINGEKYCGERSQFVVTSNSNKITVRFHSDQSYTDTGFLAEYLSYDSS
                     DPCPGQFTCRTGRCIRKELRCDGWADCTDHSDELNCSCDAGHQFTCKNKFCKPLFWVC
                     DSVNDCGDNSDEQGCSCPAQTFRCSNGKCLSKSQQCNGKDDCGDGSDEASCPKVNVVT
                     CTKHTYRCLNGLCLSKGNPECDGKEDCSDGSDEKDCDCGLRSFTRQARVVGGTDADEG
                     EWPWQVSLHALGQGHICGASLISPNWLVSAAHCYIDDRGFRYSDPTQWTVFLGLHDQS
                     QRSAPGVQERRLKRIISHPFFNDFTFDYDIALLELEKPAEYSSMVRPICLPDASHVFP
                     AGKAIWVTGWGHTQYGGTGALILQKGEIRVINQTTCENLLPQQITPRMMCVGFLSGGV
                     DSCQGDSGGPLSSVEADGRIFQAGVVSWGDGCAQRNKPGVYTRLPLFRDWIKENTGV"
ORIGIN
```

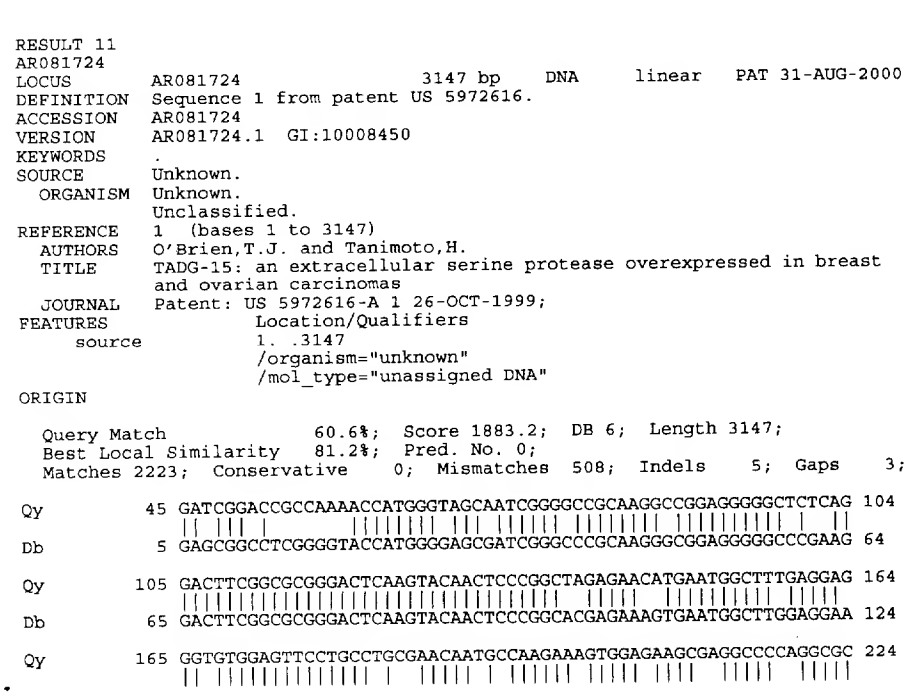
```
  Query Match           60.7%;  Score 1884.6;  DB 9;  Length 3128;
  Best Local Similarity  81.4%;  Pred. No. 0;
  Matches 2222;  Conservative    0;  Mismatches  504;  Indels    5;  Gaps    3;

Qy         50 GACCGCCAAAACCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAGGACTT 109
```

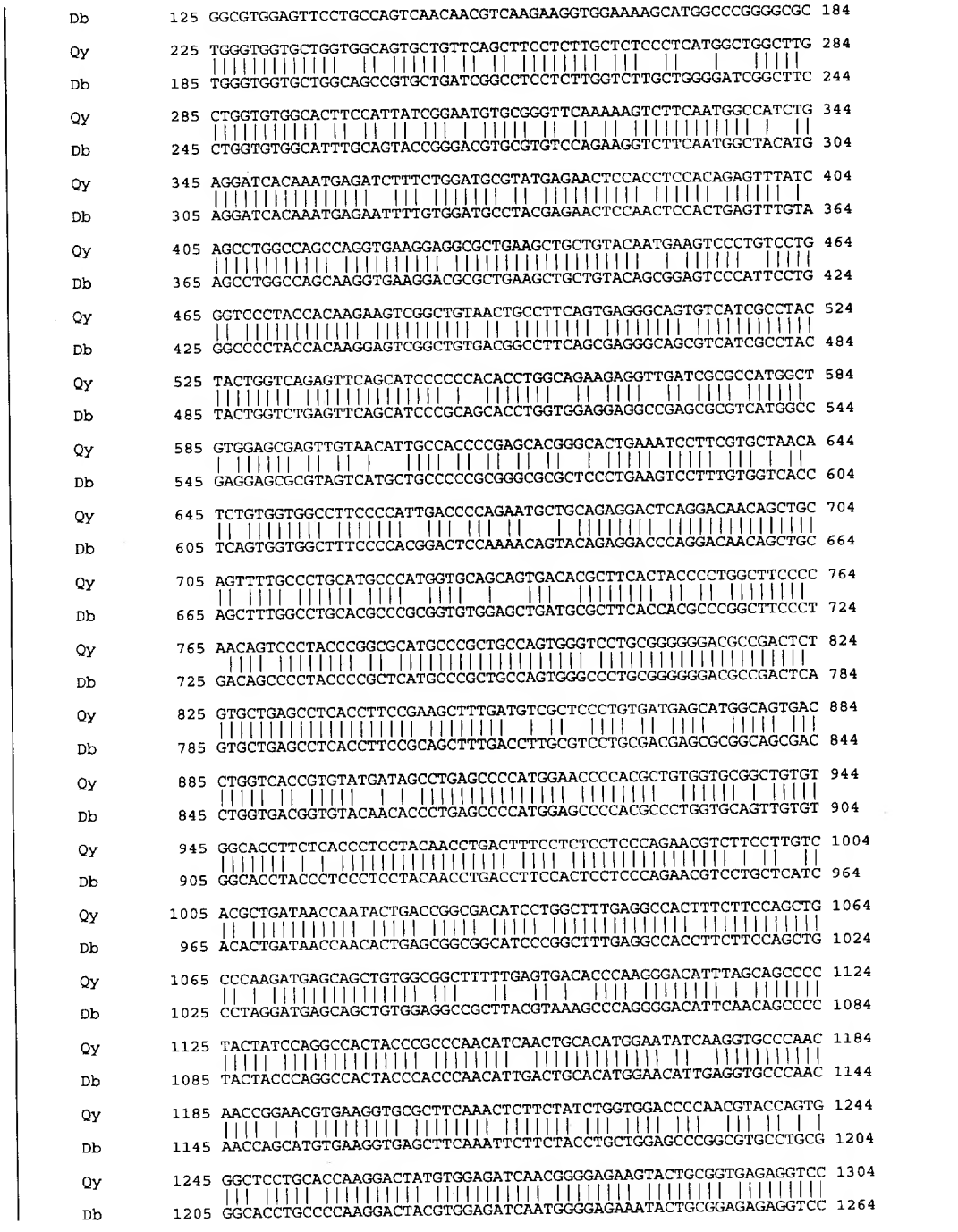

Db 3 GGCCTCGGGGACCATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGGCCCGAAGGACTT 62
Qy 110 CGGCGCGGGACTCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAGGGTGT 169
Db 63 CGGCGCGGGACTCAAGTACAACTCCCGGCACGAGAAAGTGAATGGCTTGGAGGAAGGCGT 122
Qy 170 GGAGTTCCTGCCTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGCTGGGT 229
Db 123 GGAGTTCCTGCCAGTCAACAACGTCAAGAAGGTGGAAAAGCATGGCCCGGGGCGCTGGGT 182
Qy 230 GGTGCTGGTGGCAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTGCTGGT 289
Db 183 GGTGCTGGCAGCCGTGCTGATCGGCCTCCTCTTGGTCTTGCTGGGGATCGGCTTCCTGGT 242
Qy 290 GTGGCACTTCCATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTGAGGAT 349
Db 243 GTGGCATTTGCAGTACCGGGACGTGCGTGTCCAGAAGGTCTTCAATGGCTACATGAGGAT 302
Qy 350 CACAAATGAGATCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATCAGCCT 409
Db 303 CACAAATGAGAATTTTGTGGATGCCTACGAGAACTCCAACTCCACTGAGTTTGTAAGCCT 362
Qy 410 GGCCAGCCAGGTGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTGGGTCC 469
Db 363 GGCCAGCAAGGTGAAGGACGCGCTGAAGCTGCTGTACAGCGGAGTCCCATTCCTGGGCCC 422
Qy 470 CTACCACAAGAAGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTACTACTG 529
Db 423 CTACCACAAGGAGTCGGCTGTGACGGCCTTCAGCGAGGGCAGCGTCATCGCCTACTACTG 482
Qy 530 GTCAGAGTTCAGCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCTGTGGA 589
Db 483 GTCTGAGTTCAGCATCCCGCAGCACCTGGTGGAGGAGGCCGAGCGCGTCATGGCCGAGGA 542
Qy 590 GCGAGTTGTAACATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACATCTGT 649
Db 543 GCGCGTAGTCATGCTGCCCCCGCGGGCGCGCTCCCTGAAGTCCTTTGTGGTCACCTCAGT 602
Qy 650 GGTGGCCTTCCCCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGCAGTTT 709
Db 603 GGTGGCTTTCCCCACGGACTCCAAAACAGTACAGAGGACCCAGGACAACAGCTGCAGCTT 662
Qy 710 TGCCCTGCATGCCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCCAACAG 769
Db 663 TGGCCTGCACGCCCGCGGTGTGGAGCTGATGCGCTTCACCACGCCCGGCTTCCCTGACAG 722
Qy 770 TCCCTACCCGGCGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCTGTGCT 829
Db 723 CCCCTACCCCGCTCATGCCCGCTGCCAGTGGGCCCTGCGGGGGGACGCCGACTCAGTGCT 782
Qy 830 GAGCCTCACCTTCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGACCTGGT 889
Db 783 GAGCCTCACCTTCCGCAGCTTTGACCTTGCGTCCTGCGACGAGCGCGGCAGCGACCTGGT 842
Qy 890 CACCGTGTATGATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGTGGCAC 949
Db 843 GACGGTGTACAACACCCTGAGCCCCATGGAGCCCCACGCCCTGGTGCAGTTGTGTGGCAC 902
Qy 950 CTTCTCACCCTCCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTCACGCT 1009
Db 903 CTACCCTCCCTCCTACAACCTGACCTTCCACTCCTCCCAGAACGTCCTGCTCATCACACT 962
Qy 1010 GATAACCAATACTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTGCCCAA 1069
Db 963 GATAACCAACACTGAGCGGCGGCATCCCGGCTTTGAGGCCACCTTCTTCCAGCTGCCTAG 1022
Qy 1070 GATGAGCAGCTGTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCCTACTA 1129
Db 1023 GATGAGCAGCTGTGGAGGCCGCTTACGTAAAGCCCAGGGGACATTCAACAGCCCCTACTA 1082
Qy 1130 TCCAGGCCACTACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAACAACCG 1189
Db 1083 CCCAGGCCACTACCCACCCAACATTGACTGCACATGGAACATTGAGGTGCCCAACAACCA 1142
Qy 1190 GAACGTGAAGGTGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTGGGCTC 1249
Db 1143 GCATGTGAAGGTGCGCTTCAAATTCTTCTACCTGCTGGAGCCCGGCGTGCCTGCGGGCAC 1202
Qy 1250 CTGCACCAAGGACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCCCAGTT 1309
Db 1203 CTGCCCCAAGGACTACGTGGAGATCAACGGGGAGAAATACTGCGGAGAGAGGTCCCAGTT 1262
Qy 1310 TGTGGTGAGCAGCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCGTACAC 1369
Db 1263 CGTCGTCACCAGCAACAGCAACAAGATCACAGTTCGCTTCCACTCAGATCAGTCCTACAC 1322
Qy 1370 GGACACCGGGTTCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCAGGGAT 1429
Db 1323 CGACACCGGCTTCTTAGCTGAATACCTCTCCTACGACTCCAGTGACCCATGCCCGGGGCA 1382
Qy 1430 GTTCATGTGCAAGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGGGCAGA 1489
Db 1383 GTTCACGTGCCGCACGGGGCGGTGTATCCGGAAGGAGCTGCGCTGTGATGGCTGGGCCGA 1442
Qy 1490 CTGCCCGGATTATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTCACGTG 1549
Db 1443 CTGCACCGACCACAGCGATGAGCTCAACTGCAGTTGCGACGCCGGCCACCAGTTCACGTG 1502
Qy 1550 CAAAAACCAGTTCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGTGGGGA 1609
Db 1503 CAAGAACAAGTTCTGCAAGCCCCTCTTCTGGGTCTGCGACAGTGTGAACGACTGCGGAGA 1562
Qy 1610 CGGAAGTGACGAGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAATGGGAA 1669
Db 1563 CAACAGCGACGAGCAGGGGTGCAGTTGTCCGGCCCAGACCTTCAGGTGTTCCAATGGGAA 1622
Qy 1670 GTGTCTCCCTCAGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCTGACGA 1729
Db 1623 GTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTGGGGACGGGTCCGACGA 1682
Qy 1730 GGCTTCATGTGACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGCCAAAA 1789
Db 1683 GGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCTACCGCTGCCTCAA 1742
Qy 1790 TGGCCTCTGTCTGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGCGATGG 1849
Db 1743 TGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGGAGGACTGTAGCGACGG 1802
Qy 1850 CTCCGATGAGAAAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGCGTGGT 1909
Db 1803 CTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGTCATTCACGAGACAGGCTCGTGTTGT 1862
Qy 1910 TGGTGGCACGAATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCCCTGGG 1969
Db 1863 TGGGGGCACGGATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAAGCCTGCATGCTCTGGG 1922
Qy 1970 CCAGGGCCACTTGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCAGCTCA 2029
Db 1923 CCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGCTGGTCTCTGCCGCACA 1982
Qy 2030 TTGCTTTCAGGATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCCTTCCT 2089
Db 1983 CTGCTACATCGATGACAGAGGATTCAGGTACTCAGACCCCACGCAGTGGACGGTCTTCCT 2042
Qy 2090 GGGTCTGCTGGACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTCAAACG 2149
Db 2043 GGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGGAGCGCAGGCTCAAGCG 2102
Qy 2150 TATCATCACCCACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTGCTGGA 2209
Db 2103 CATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATGACATCGCGCTGCTGGA 2162
Qy 2210 GCTGGAGAAGTCGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGATGCTAC 2269
Db 2163 GCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGACGCCTC 2222



RESULT 11
AR081724
LOCUS AR081724 3147 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 1 from patent US 5972616.
ACCESSION AR081724
VERSION AR081724.1 GI:10008450
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 3147)
AUTHORS O'Brien,T.J. and Tanimoto,H.
TITLE TADG-15: an extracellular serine protease overexpressed in breast and ovarian carcinomas
JOURNAL Patent: US 5972616-A 1 26-OCT-1999;
FEATURES Location/Qualifiers
source 1. .3147
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 60.6%; Score 1883.2; DB 6; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;



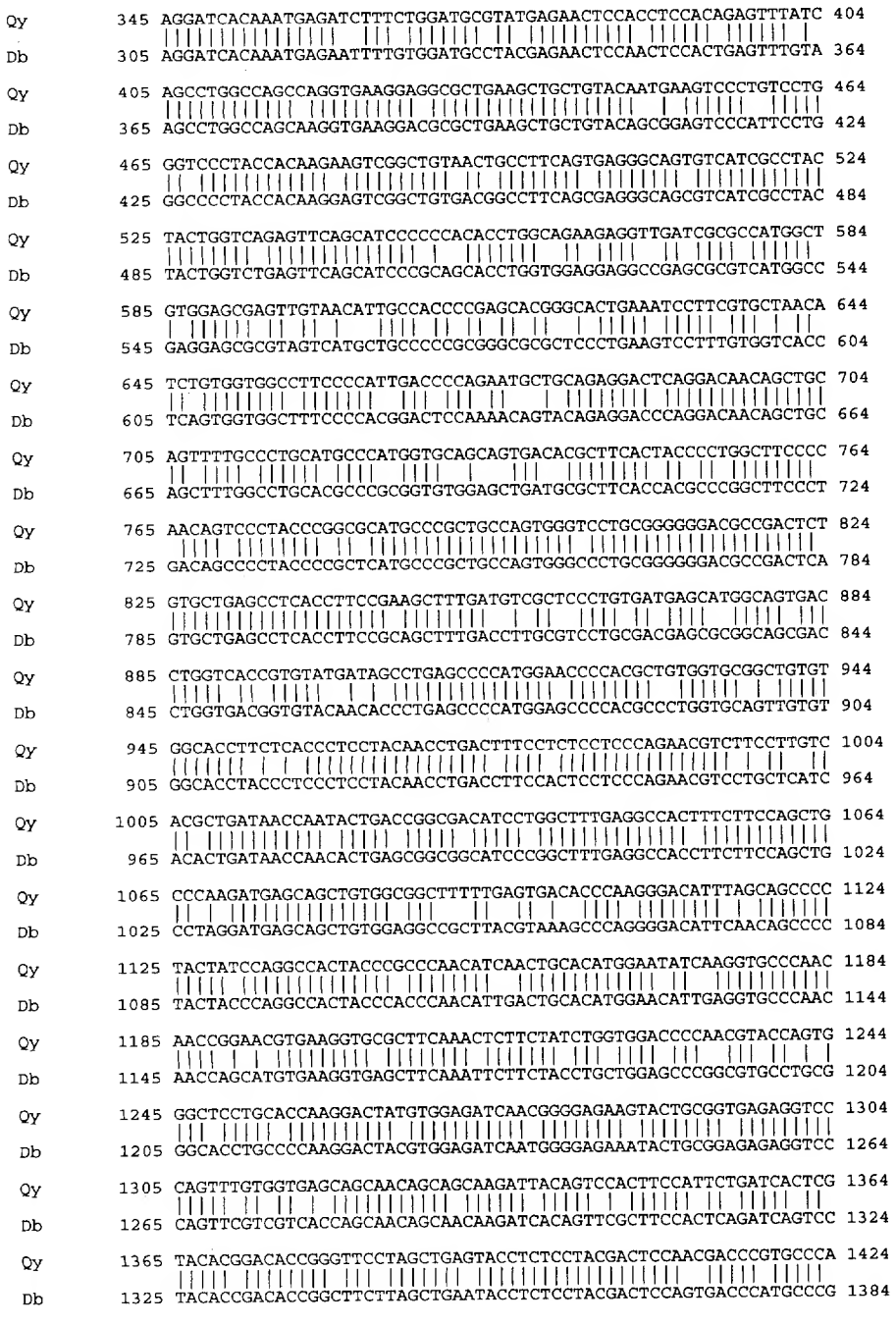


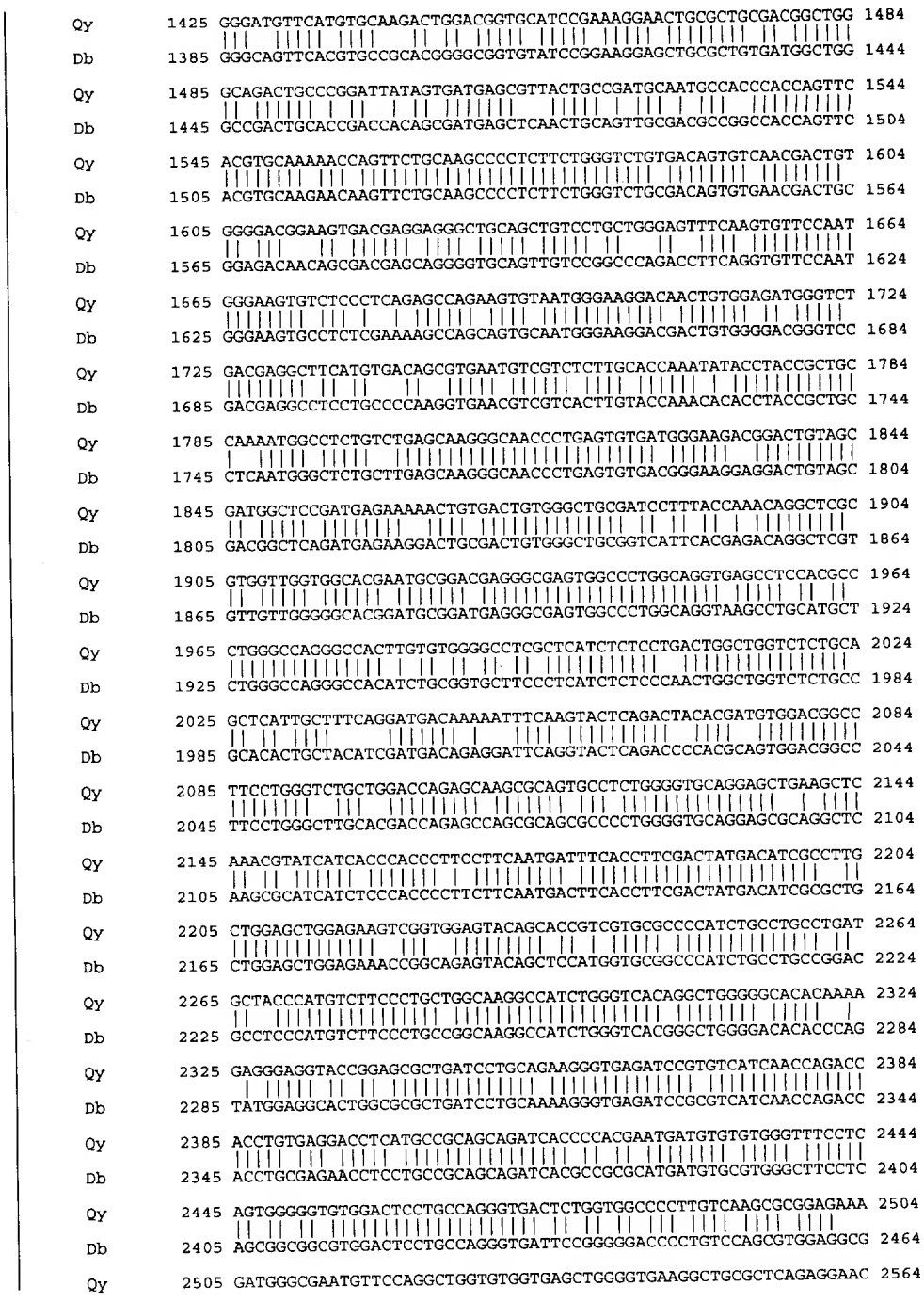
```
Qy       1305 CAGTTTGTGGTGAGCAGCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCG 1364
Db       1265 CAGTTCGTCGTCACCAGCAACAGCAACAAGATCACAGTTCGCTTCCACTCAGATCAGTCC 1324

Qy       1365 TACACGGACACCGGGTTCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCA 1424
Db       1325 TACACCGACACCGGCTTCTTAGCTGAATACCTCTCCTACGACTCCAGTGACCCATGCCCG 1384

Qy       1425 GGGATGTTCATGTGCAAGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGG 1484
Db       1385 GGGCAGTTCACGTGCCGCACGGGGCGGTGTATCCGGAAGGAGCTGCGCTGTGATGGCTGG 1444

Qy       1485 GCAGACTGCCCGGATTATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTC 1544
Db       1445 GCCGACTGCACCGACCACAGCGATGAGCTCAACTGCAGTTGCGACGCCGGCCACCAGTTC 1504

Qy       1545 ACGTGCAAAAACCAGTTCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGT 1604
Db       1505 ACGTGCAAGAACAAGTTCTGCAAGCCCCTCTTCTGGGTCTGCGACAGTGTGAACGACTGC 1564

Qy       1605 GGGGACGGAAGTGACGAGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAAT 1664
Db       1565 GGAGACAACAGCGACGAGCAGGGGTGCAGTTGTCCGGCCCAGACCTTCAGGTGTTCCAAT 1624

Qy       1665 GGGAAGTGTCTCCCTCAGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCT 1724
Db       1625 GGGAAGTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTGGGGACGGGTCC 1684

Qy       1725 GACGAGGCTTCATGTGACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGC 1784
Db       1685 GACGAGGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCTACCGCTGC 1744

Qy       1785 CAAAATGGCCTCTGTCTGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGC 1844
Db       1745 CTCAATGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGGAGGACTGTAGC 1804

Qy       1845 GATGGCTCCGATGAGAAAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGC 1904
Db       1805 GACGGCTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGTCATTCACGAGACAGGCTCGT 1864

Qy       1905 GTGGTTGGTGGCACGAATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCC 1964
Db       1865 GTTGTTGGGGGCACGGATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAAGCCTGCATGCT 1924

Qy       1965 CTGGGCCAGGGCCACTTGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCA 2024
Db       1925 CTGGGCCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGCTGGTCTCTGCC 1984

Qy       2025 GCTCATTGCTTTCAGGATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCC 2084
Db       1985 GCACACTGCTACATCGATGACAGAGGATTCAGGTACTCAGACCCCACGCAGTGGACGGCC 2044

Qy       2085 TTCCTGGGTCTGCTGGACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTC 2144
Db       2045 TTCCTGGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGGAGCGCAGGCTC 2104

Qy       2145 AAACGTATCATCACCCACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTG 2204
Db       2105 AAGCGCATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATGACATCGCGCTG 2164

Qy       2205 CTGGAGCTGGAGAAGTCGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGAT 2264
Db       2165 CTGGAGCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGAC 2224

Qy       2265 GCTACCCATGTCTTCCCTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAA 2324
Db       2225 GCCTCCCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGGGCTGGGGACACACCCAG 2284

Qy       2325 GAGGGAGGTACCGGAGCGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACC 2384
Db       2285 TATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACC 2344

Qy       2385 ACCTGTGAGGACCTCATGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTC 2444
Db       2345 ACCTGCGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGATGTGCGTGGGCTTCCTC 2404

Qy       2445 AGTGGGGGTGTGGACTCCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAA 2504
Db       2405 AGCGGCGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGTGGAGGCG 2464

Qy       2505 GATGGGCGAATGTTCCAGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAAC 2564
Db       2465 GATGGGCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAAC 2524

Qy       2565 AAGCCAGGCGTGTACACAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGG 2624
Db       2525 AAGCCAGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAGAGAACACTGGG 2584

Qy       2625 GTATAGCAGCATGGACAGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACA 2684
Db       2585 GTATAGGGGCCGGG---GCCACCCAAATGTGTACACCTGCGGGGCCACCCATCGTCCACC 2641

Qy       2685 CCTGGATACAGGAGAGGAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACA 2744
Db       2642 CCAGTGTGCACG-CCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCC-CCAGA 2699

Qy       2745 ACCCAGACTGTGAACTGCATCCTTAGGACTCAGAGT 2780
Db       2700 ACATACACTGTGAACTCAATCTCCAGGGCTCCAAAT 2735
```

RESULT 12
AR229704
LOCUS AR229704 3147 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6451500.
ACCESSION AR229704
VERSION AR229704.1 GI:27269434
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 3147)
AUTHORS Leon,J.W.
TITLE Imaging member containing heat switchable carboxylate polymer and method of use
JOURNAL Patent: US 6451500-A 1 17-SEP-2002;
FEATURES Location/Qualifiers
source 1. .3147
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 60.6%; Score 1883.2; DB 6; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

```
Qy         45 GATCGGACCGCCAAAACCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAG 104
Db          5 GAGCGGCCTCGGGGTACCATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGGCCCGAAG 64

Qy        105 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAG 164
Db         65 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCACGAGAAAGTGAATGGCTTGGAGGAA 124

Qy        165 GGTGTGGAGTTCCTGCCTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGC 224
Db        125 GGCGTGGAGTTCCTGCCAGTCAACAACGTCAAGAAGGTGGAAAAGCATGGCCCGGGGCGC 184

Qy        225 TGGGTGGTGCTGGTGGCAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTG 284
Db        185 TGGGTGGTGCTGGCAGCCGTGCTGATCGGCCTCCTCTTGGTCTTGCTGGGGATCGGCTTC 244

Qy        285 CTGGTGTGGCACTTCCATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTG 344
Db        245 CTGGTGTGGCATTTGCAGTACCGGGACGTGCGTGTCCAGAAGGTCTTCAATGGCTACATG 304
```





```
Db       2465 GATGGGCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAAC 2524

Qy       2565 AAGCCAGGCGTGTACACAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGG 2624

Db       2525 AAGCCAGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAGAGAACACTGGG 2584

Qy       2625 GTATAGCAGCATGGACAGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACA 2684

Db       2585 GTATAGGGGCCGGG---GCCACCCAAATGTGTACACCTGCGGGGCCACCCATCGTCCACC 2641

Qy       2685 CCTGGATACAGGAGAGGAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACA 2744

Db       2642 CCAGTGTGCACG-CCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCC-CCAGA 2699

Qy       2745 ACCCAGACTGTGAACTGCATCCTTAGGACTCAGAGT 2780

Db       2700 ACATACACTGTGAACTCAATCTCCAGGGCTCCAAAT 2735


RESULT 13
AR229712/c
LOCUS       AR229712                3147 bp    RNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 18 from patent US 6451500.
ACCESSION   AR229712
VERSION     AR229712.1  GI:27269442
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3147)
  AUTHORS   Leon,J.W.
  TITLE     Imaging member containing heat switchable carboxylate polymer and
            method of use
  JOURNAL   Patent: US 6451500-A 18 17-SEP-2002;
FEATURES             Location/Qualifiers
     source          1. .3147
                     /organism="unknown"
                     /mol_type="unassigned RNA"
ORIGIN

  Query Match           60.6%;  Score 1883.2;  DB 6;  Length 3147;
  Best Local Similarity   81.2%;  Pred. No. 0;
  Matches 2223;  Conservative    0;  Mismatches  508;  Indels    5;  Gaps    3;

Qy         45 GATCGGACCGCCAAAACCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAG 104

Db       3143 GAGCGGCCTCGGGGTACCATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGGCCCGAAG 3084

Qy        105 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAG 164

Db       3083 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCACGAGAAAGTGAATGGCTTGGAGGAA 3024

Qy        165 GGTGTGGAGTTCCTGCCTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGC 224

Db       3023 GGCGTGGAGTTCCTGCCAGTCAACAACGTCAAGAAGGTGGAAAAGCATGGCCCGGGGCGC 2964

Qy        225 TGGGTGGTGCTGGTGGCAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTG 284

Db       2963 TGGGTGGTGCTGGCAGCCGTGCTGATCGGCCTCCTCTTGGTCTTGCTGGGGATCGGCTTC 2904

Qy        285 CTGGTGTGGCACTTCCATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTG 344

Db       2903 CTGGTGTGGCATTTGCAGTACCGGGACGTGCGTGTCCAGAAGGTCTTCAATGGCTACATG 2844

Qy        345 AGGATCACAAATGAGATCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATC 404

Db       2843 AGGATCACAAATGAGAATTTTGTGGATGCCTACGAGAACTCCAACTCCACTGAGTTTGTA 2784

Qy        405 AGCCTGGCCAGCCAGGTGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTG 464

Db       2783 AGCCTGGCCAGCAAGGTGAAGGACGCGCTGAAGCTGCTGTACAGCGGAGTCCCATTCCTG 2724

Qy        465 GGTCCCTACCACAAGAAGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTAC 524

Db       2723 GGCCCCTACCACAAGGAGTCGGCTGTGACGGCCTTCAGCGAGGGCAGCGTCATCGCCTAC 2664

Qy        525 TACTGGTCAGAGTTCAGCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCT 584

Db       2663 TACTGGTCTGAGTTCAGCATCCCGCAGCACCTGGTGGAGGAGGCCGAGCGCGTCATGGCC 2604

Qy        585 GTGGAGCGAGTTGTAACATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACA 644

Db       2603 GAGGAGCGCGTAGTCATGCTGCCCCCGCGGGCGCGCTCCCTGAAGTCCTTTGTGGTCACC 2544

Qy        645 TCTGTGGTGGCCTTCCCCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGC 704

Db       2543 TCAGTGGTGGCTTTCCCCACGGACTCCAAAACAGTACAGAGGACCCAGGACAACAGCTGC 2484

Qy        705 AGTTTTGCCCTGCATGCCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCC 764

Db       2483 AGCTTTGGCCTGCACGCCCGCGGTGTGGAGCTGATGCGCTTCACCACGCCCGGCTTCCCT 2424

Qy        765 AACAGTCCCTACCCGGCGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCT 824

Db       2423 GACAGCCCCTACCCCGCTCATGCCCGCTGCCAGTGGGCCCTGCGGGGGGACGCCGACTCA 2364

Qy        825 GTGCTGAGCCTCACCTTCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGAC 884

Db       2363 GTGCTGAGCCTCACCTTCCGCAGCTTTGACCTTGCGTCCTGCGACGAGCGCGGCAGCGAC 2304

Qy        885 CTGGTCACCGTGTATGATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGT 944

Db       2303 CTGGTGACGGTGTACAACACCCTGAGCCCCATGGAGCCCCACGCCCTGGTGCAGTTGTGT 2244

Qy        945 GGCACCTTCTCACCCTCCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTC 1004

Db       2243 GGCACCTACCCTCCCTCCTACAACCTGACCTTCCACTCCTCCCAGAACGTCCTGCTCATC 2184

Qy       1005 ACGCTGATAACCAATACTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTG 1064

Db       2183 ACACTGATAACCAACACTGAGCGGCGGCATCCCGGCTTTGAGGCCACCTTCTTCCAGCTG 2124

Qy       1065 CCCAAGATGAGCAGCTGTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCC 1124

Db       2123 CCTAGGATGAGCAGCTGTGGAGGCCGCTTACGTAAAGCCCAGGGGACATTCAACAGCCCC 2064

Qy       1125 TACTATCCAGGCCACTACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAAC 1184

Db       2063 TACTACCCAGGCCACTACCCACCCAACATTGACTGCACATGGAACATTGAGGTGCCCAAC 2004

Qy       1185 AACCGGAACGTGAAGGTGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTG 1244

Db       2003 AACCAGCATGTGAAGGTGAGCTTCAAATTCTTCTACCTGCTGGAGCCCGGCGTGCCTGCG 1944

Qy       1245 GGCTCCTGCACCAAGGACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCC 1304

Db       1943 GGCACCTGCCCCAAGGACTACGTGGAGATCAATGGGGAGAAATACTGCGGAGAGAGGTCC 1884

Qy       1305 CAGTTTGTGGTGAGCAGCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCG 1364

Db       1883 CAGTTCGTCGTCACCAGCAACAGCAACAAGATCACAGTTCGCTTCCACTCAGATCAGTCC 1824

Qy       1365 TACACGGACACCGGGTTCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCA 1424

Db       1823 TACACCGACACCGGCTTCTTAGCTGAATACCTCTCCTACGACTCCAGTGACCCATGCCCG 1764

Qy       1425 GGGATGTTCATGTGCAAGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGG 1484

Db       1763 GGGCAGTTCACGTGCCGCACGGGGCGGTGTATCCGGAAGGAGCTGCGCTGTGATGGCTGG 1704

Qy       1485 GCAGACTGCCCGGATTATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTC 1544

Db       1703 GCCGACTGCACCGACCACAGCGATGAGCTCAACTGCAGTTGCGACGCCGGCCACCAGTTC 1644

Qy       1545 ACGTGCAAAAACCAGTTCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGT 1604
```

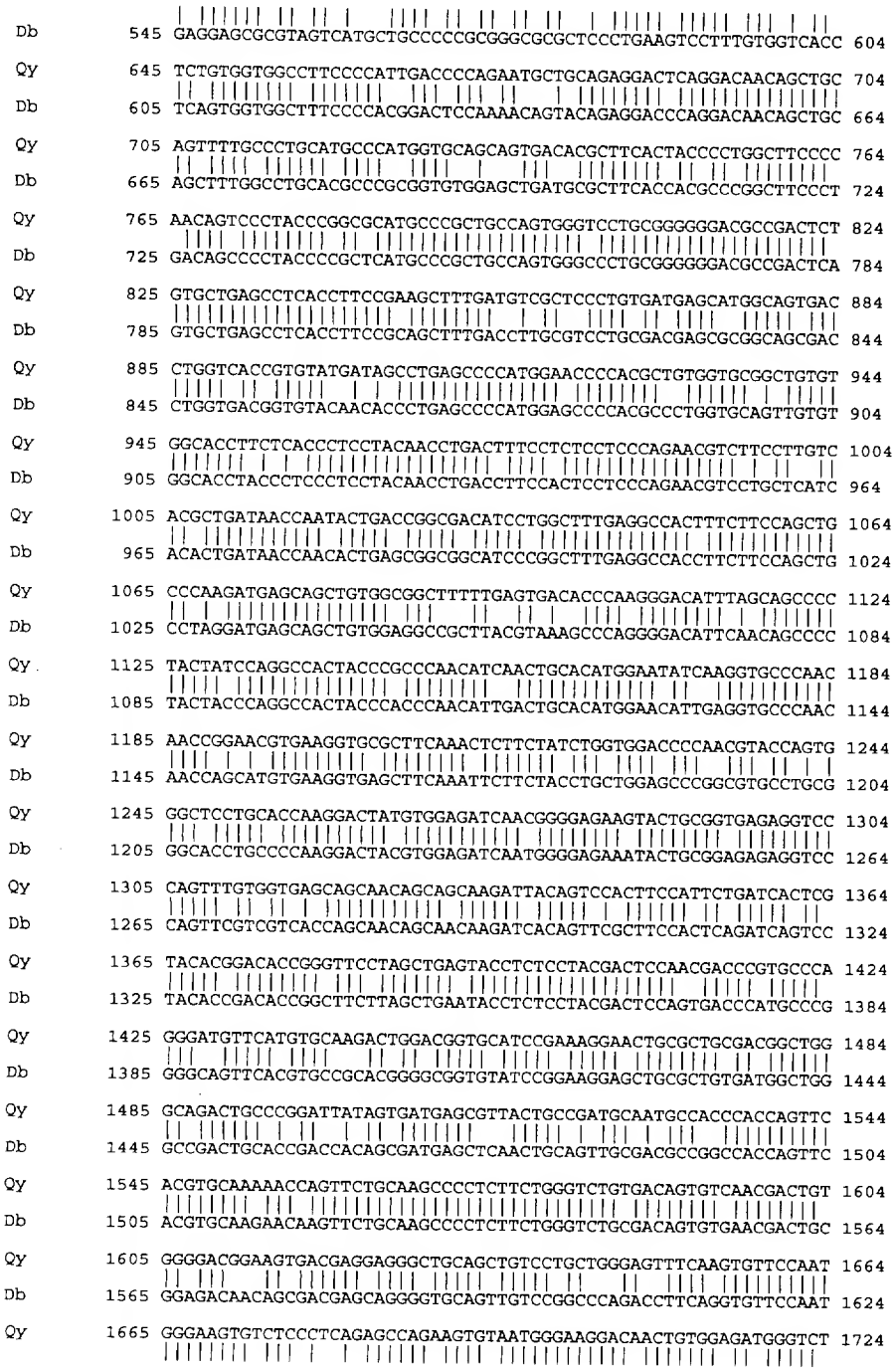
```
Db      1643 ACGTGCAAGAACAAGTTCTGCAAGCCCCTCTTCTGGGTCTGCGACAGTGTGAACGACTGC 1584
Qy      1605 GGGGACGGAAGTGACGAGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAAT 1664
Db      1583 GGAGACAACAGCGACGAGCAGGGGTGCAGTTGTCCGGCCCAGACCTTCAGGTGTTCCAAT 1524
Qy      1665 GGGAAGTGTCTCCCTCAGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCT 1724
Db      1523 GGGAAGTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTGGGGACGGGTCC 1464
Qy      1725 GACGAGGCTTCATGTGACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGC 1784
Db      1463 GACGAGGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCTACCGCTGC 1404
Qy      1785 CAAAATGGCCTCTGTCTGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGC 1844
Db      1403 CTCAATGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGGAGGACTGTAGC 1344
Qy      1845 GATGGCTCCGATGAGAAAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGC 1904
Db      1343 GACGGCTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGTCATTCACGAGACAGGCTCGT 1284
Qy      1905 GTGGTTGGTGGCACGAATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCC 1964
Db      1283 GTTGTTGGGGGCACGGATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAAGCCTGCATGCT 1224
Qy      1965 CTGGGCCAGGGCCACTTGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCA 2024
Db      1223 CTGGGCCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGCTGGTCTCTGCC 1164
Qy      2025 GCTCATTGCTTTCAGGATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCC 2084
Db      1163 GCACACTGCTACATCGATGACAGAGGATTCAGGTACTCAGACCCCACGCAGTGGACGGCC 1104
Qy      2085 TTCCTGGGTCTGCTGGACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTC 2144
Db      1103 TTCCTGGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGGAGCGCAGGCTC 1044
Qy      2145 AAACGTATCATCACCCACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTG 2204
Db      1043 AAGCGCATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATGACATCGCGCTG 984
Qy      2205 CTGGAGCTGGAGAAGTCGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGAT 2264
Db       983 CTGGAGCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGAC 924
Qy      2265 GCTACCCATGTCTTCCCTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAA 2324
Db       923 GCCTCCCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGGGCTGGGGACACACCCAG 864
Qy      2325 GAGGGAGGTACCGGAGCGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACC 2384
Db       863 TATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACC 804
Qy      2385 ACCTGTGAGGACCTCATGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTC 2444
Db       803 ACCTGCGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGATGTGCGTGGGCTTCCTC 744
Qy      2445 AGTGGGGGTGTGGACTCCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAA 2504
Db       743 AGCGGCGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGTGGAGGCG 684
Qy      2505 GATGGGCGAATGTTCCAGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAAC 2564
Db       683 GATGGGCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAAC 624
Qy      2565 AAGCCAGGCGTGTACACAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGG 2624
Db       623 AAGCCAGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAGAGAACACTGGG 564
Qy      2625 GTATAGCAGCATGGACAGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACA 2684
```

```
Db       563 GTATAGGGGCCGGG---GCCACCCAAATGTGTACACCTGCGGGGCCACCCATCGTCCACC 507
Qy      2685 CCTGGATACAGGAGAGGAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACA 2744
Db       506 CCAGTGTGCACG-CCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCC-CCAGA 449
Qy      2745 ACCCAGACTGTGAACTGCATCCTTAGGACTCAGAGT 2780
Db       448 ACATACACTGTGAACTCAATCTCCAGGGCTCCAAAT 413
```

RESULT 14
AR430930
LOCUS AR430930 3147 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6649741.
ACCESSION AR430930
VERSION AR430930.1 GI:40192672
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 3147)
AUTHORS O'Brien,T.J. and Tanimoto,H.
TITLE TADG-15: an extracellular serine protease overexpressed in carcinomas
JOURNAL Patent: US 6649741-A 1 18-NOV-2003;
FEATURES Location/Qualifiers
source 1. .3147
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 60.6%; Score 1883.2; DB 6; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

```
Qy        45 GATCGGACCGCCAAAACCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAG 104
Db         5 GAGCGGCCTCGGGGTACCATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGGCCCGAAG 64
Qy       105 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAG 164
Db        65 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCACGAGAAAGTGAATGGCTTGGAGGAA 124
Qy       165 GGTGTGGAGTTCCTGCCTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGC 224
Db       125 GGCGTGGAGTTCCTGCCAGTCAACAACGTCAAGAAGGTGGAAAAGCATGGCCCGGGGCGC 184
Qy       225 TGGGTGGTGCTGGTGGCAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTG 284
Db       185 TGGGTGGTGCTGGCAGCCGTGCTGATCGGCCTCCTCTTGGTCTTGCTGGGGATCGGCTTC 244
Qy       285 CTGGTGTGGCACTTCCATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTG 344
Db       245 CTGGTGTGGCATTTGCAGTACCGGGACGTGCGTGTCCAGAAGGTCTTCAATGGCTACATG 304
Qy       345 AGGATCACAAATGAGATCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATC 404
Db       305 AGGATCACAAATGAGAATTTTGTGGATGCCTACGAGAACTCCAACTCCACTGAGTTTGTA 364
Qy       405 AGCCTGGCCAGCCAGGTGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTG 464
Db       365 AGCCTGGCCAGCAAGGTGAAGGACGCGCTGAAGCTGCTGTACAGCGGAGTCCCATTCCTG 424
Qy       465 GGTCCCTACCACAAGAAGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTAC 524
Db       425 GGCCCCTACCACAAGGAGTCGGCTGTGACGGCCTTCAGCGAGGGCAGCGTCATCGCCTAC 484
Qy       525 TACTGGTCAGAGTTCAGCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCT 584
Db       485 TACTGGTCTGAGTTCAGCATCCCGCAGCACCTGGTGGAGGAGGCCGAGCGCGTCATGGCC 544
Qy       585 GTGGAGCGAGTTGTAACATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACA 644
```


Db 545 GAGGAGCGCGTAGTCATGCTGCCCCCGCGGGCGCGCTCCCTGAAGTCCTTTGTGGTCACC 604
Qy 645 TCTGTGGTGGCCTTCCCCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGC 704
Db 605 TCAGTGGTGGCTTTCCCCACGGACTCCAAAACAGTACAGAGGACCCAGGACAACAGCTGC 664
Qy 705 AGTTTTGCCCTGCATGCCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCC 764
Db 665 AGCTTTGGCCTGCACGCCCGCGGTGTGGAGCTGATGCGCTTCACCACGCCCGGCTTCCCT 724
Qy 765 AACAGTCCCTACCCGGCGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCT 824
Db 725 GACAGCCCCTACCCCGCTCATGCCCGCTGCCAGTGGGCCCTGCGGGGGGACGCCGACTCA 784
Qy 825 GTGCTGAGCCTCACCTTCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGAC 884
Db 785 GTGCTGAGCCTCACCTTCCGCAGCTTTGACCTTGCGTCCTGCGACGAGCGCGGCAGCGAC 844
Qy 885 CTGGTCACCGTGTATGATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGT 944
Db 845 CTGGTGACGGTGTACAACACCCTGAGCCCCATGGAGCCCCACGCCCTGGTGCAGTTGTGT 904
Qy 945 GGCACCTTCTCACCCTCCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTC 1004
Db 905 GGCACCTACCCTCCCTCCTACAACCTGACCTTCCACTCCTCCCAGAACGTCCTGCTCATC 964
Qy 1005 ACGCTGATAACCAATACTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTG 1064
Db 965 ACACTGATAACCAACACTGAGCGGCGGCATCCCGGCTTTGAGGCCACCTTCTTCCAGCTG 1024
Qy 1065 CCCAAGATGAGCAGCTGTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCC 1124
Db 1025 CCTAGGATGAGCAGCTGTGGAGGCCGCTTACGTAAAGCCCAGGGGACATTCAACAGCCCC 1084
Qy 1125 TACTATCCAGGCCACTACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAAC 1184
Db 1085 TACTACCCAGGCCACTACCCACCCAACATTGACTGCACATGGAACATTGAGGTGCCCAAC 1144
Qy 1185 AACCGGAACGTGAAGGTGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTG 1244
Db 1145 AACCAGCATGTGAAGGTGAGCTTCAAATTCTTCTACCTGCTGGAGCCCGGCGTGCCTGCG 1204
Qy 1245 GGCTCCTGCACCAAGGACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCC 1304
Db 1205 GGCACCTGCCCCAAGGACTACGTGGAGATCAATGGGGAGAAATACTGCGGAGAGAGGTCC 1264
Qy 1305 CAGTTTGTGGTGAGCAGCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCG 1364
Db 1265 CAGTTCGTCGTCACCAGCAACAGCAACAAGATCACAGTTCGCTTCCACTCAGATCAGTCC 1324
Qy 1365 TACACGGACACCGGGTTCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCA 1424
Db 1325 TACACCGACACCGGCTTCTTAGCTGAATACCTCTCCTACGACTCCAGTGACCCATGCCCG 1384
Qy 1425 GGGATGTTCATGTGCAAGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGG 1484
Db 1385 GGGCAGTTCACGTGCCGCACGGGGCGGTGTATCCGGAAGGAGCTGCGCTGTGATGGCTGG 1444
Qy 1485 GCAGACTGCCCGGATTATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTC 1544
Db 1445 GCCGACTGCACCGACCACAGCGATGAGCTCAACTGCAGTTGCGACGCCGGCCACCAGTTC 1504
Qy 1545 ACGTGCAAAAACCAGTTCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGT 1604
Db 1505 ACGTGCAAGAACAAGTTCTGCAAGCCCCTCTTCTGGGTCTGCGACAGTGTGAACGACTGC 1564
Qy 1605 GGGGACGGAAGTGACGAGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAAT 1664
Db 1565 GGAGACAACAGCGACGAGCAGGGGTGCAGTTGTCCGGCCCAGACCTTCAGGTGTTCCAAT 1624
Qy 1665 GGGAAGTGTCTCCCTCAGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCT 1724
Db 1625 GGGAAGTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTGGGGACGGGTCC 1684
Qy 1725 GACGAGGCTTCATGTGACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGC 1784
Db 1685 GACGAGGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCTACCGCTGC 1744
Qy 1785 CAAAATGGCCTCTGTCTGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGC 1844
Db 1745 CTCAATGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGGAGGACTGTAGC 1804
Qy 1845 GATGGCTCCGATGAGAAAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGC 1904
Db 1805 GACGGCTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGTCATTCACGAGACAGGCTCGT 1864
Qy 1905 GTGGTTGGTGGCACGAATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCC 1964
Db 1865 GTTGTTGGGGGCACGGATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAAGCCTGCATGCT 1924
Qy 1965 CTGGGCCAGGGCCACTTGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCA 2024
Db 1925 CTGGGCCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGCTGGTCTCTGCC 1984
Qy 2025 GCTCATTGCTTTCAGGATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCC 2084
Db 1985 GCACACTGCTACATCGATGACAGAGGATTCAGGTACTCAGACCCCACGCAGTGGACGGCC 2044
Qy 2085 TTCCTGGGTCTGCTGGACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTC 2144
Db 2045 TTCCTGGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGGAGCGCAGGCTC 2104
Qy 2145 AAACGTATCATCACCCACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTG 2204
Db 2105 AAGCGCATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATGACATCGCGCTG 2164
Qy 2205 CTGGAGCTGGAGAAGTCGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGAT 2264
Db 2165 CTGGAGCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGAC 2224
Qy 2265 GCTACCCATGTCTTCCCTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAA 2324
Db 2225 GCCTCCCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGGGCTGGGGACACACCCAG 2284
Qy 2325 GAGGGAGGTACCGGAGCGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACC 2384
Db 2285 TATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACC 2344
Qy 2385 ACCTGTGAGGACCTCATGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTC 2444
Db 2345 ACCTGCGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGATGTGCGTGGGCTTCCTC 2404
Qy 2445 AGTGGGGGTGTGGACTCCTGCCAGGGTGACTCTGGTGGCCCCCTTGTCAAGCGCGGAGAAA 2504
Db 2405 AGCGGCGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGTGGAGGCG 2464
Qy 2505 GATGGGCGAATGTTCCAGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAAC 2564
Db 2465 GATGGGCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAAC 2524
Qy 2565 AAGCCAGGCGTGTACACAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGG 2624
Db 2525 AAGCCAGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAGAGAACACTGGG 2584
Qy 2625 GTATAGCAGCATGGACAGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACA 2684
Db 2585 GTATAGGGGCCGGG---GCCACCCAAATGTGTACACCTGCGGGGCCACCCATCGTCCACC 2641
Qy 2685 CCTGGATACAGGAGAGGAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACA 2744
Db 2642 CCAGTGTGCACG-CCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCC-CCAGA 2699
Qy 2745 ACCCAGACTGTGAACTGCATCCTTAGGACTCAGAGT 2780
Db 2700 ACATACACTGTGAACTCAATCTCCAGGGCTCCAAAT 2735

```
RESULT 15
AR430938/c
LOCUS       AR430938               3147 bp    RNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 18 from patent US 6649741.
ACCESSION   AR430938
VERSION     AR430938.1  GI:40192680
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3147)
  AUTHORS   O'Brien,T.J. and Tanimoto,H.
  TITLE     TADG-15: an extracellular serine protease overexpressed in
            carcinomas
  JOURNAL   Patent: US 6649741-A 18 18-NOV-2003;
FEATURES             Location/Qualifiers
     source          1. .3147
                     /organism="unknown"
                     /mol_type="unassigned RNA"
ORIGIN

  Query Match             60.6%;  Score 1883.2;  DB 6;  Length 3147;
  Best Local Similarity   81.2%;  Pred. No. 0;
  Matches 2223;  Conservative    0;  Mismatches  508;  Indels    5;  Gaps    3;

Qy         45 GATCGGACCGCCAAAACCATGGGTAGCAATCGGGGCCGCAAGGCCGGAGGGGGCTCTCAG 104
Db       3143 GAGCGGCCTCGGGGTACCATGGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGGCCCGAAG 3084

Qy        105 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCTAGAGAACATGAATGGCTTTGAGGAG 164
Db       3083 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCACGAGAAAGTGAATGGCTTGGAGGAA 3024

Qy        165 GGTGTGGAGTTCCTGCCTGCGAACAATGCCAAGAAAGTGGAGAAGCGAGGCCCCAGGCGC 224
Db       3023 GGCGTGGAGTTCCTGCCAGTCAACAACGTCAAGAAGGTGGAAAAGCATGGCCCGGGGCGC 2964

Qy        225 TGGGTGGTGCTGGTGGCAGTGCTGTTCAGCTTCCTCTTGCTCTCCCTCATGGCTGGCTTG 284
Db       2963 TGGGTGGTGCTGGCAGCCGTGCTGATCGGCCTCCTCTTGGTCTTGCTGGGGATCGGCTTC 2904

Qy        285 CTGGTGTGGCACTTCCATTATCGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCTG 344
Db       2903 CTGGTGTGGCATTTGCAGTACCGGGACGTGCGTGTCCAGAAGGTCTTCAATGGCTACATG 2844

Qy        345 AGGATCACAAATGAGATCTTTCTGGATGCGTATGAGAACTCCACCTCCACAGAGTTTATC 404
Db       2843 AGGATCACAAATGAGAATTTTGTGGATGCCTACGAGAACTCCAACTCCACTGAGTTTGTA 2784

Qy        405 AGCCTGGCCAGCCAGGTGAAGGAGGCGCTGAAGCTGCTGTACAATGAAGTCCCTGTCCTG 464
Db       2783 AGCCTGGCCAGCAAGGTGAAGGACGCGCTGAAGCTGCTGTACAGCGGAGTCCCATTCCTG 2724

Qy        465 GGTCCCTACCACAAGAAGTCGGCTGTAACTGCCTTCAGTGAGGGCAGTGTCATCGCCTAC 524
Db       2723 GGCCCCTACCACAAGGAGTCGGCTGTGACGGCCTTCAGCGAGGGCAGCGTCATCGCCTAC 2664

Qy        525 TACTGGTCAGAGTTCAGCATCCCCCCACACCTGGCAGAAGAGGTTGATCGCGCCATGGCT 584
Db       2663 TACTGGTCTGAGTTCAGCATCCCGCAGCACCTGGTGGAGGAGGCCGAGCGCGTCATGGCC 2604

Qy        585 GTGGAGCGAGTTGTAACATTGCCACCCCGAGCACGGGCACTGAAATCCTTCGTGCTAACA 644
Db       2603 GAGGAGCGCGTAGTCATGCTGCCCCCGCGGGCGCGCTCCCTGAAGTCCTTTGTGGTCACC 2544

Qy        645 TCTGTGGTGGCCTTCCCCATTGACCCCAGAATGCTGCAGAGGACTCAGGACAACAGCTGC 704
Db       2543 TCAGTGGTGGCTTTCCCCACGGACTCCAAAACAGTACAGAGGACCCAGGACAACAGCTGC 2484

Qy        705 AGTTTTGCCCTGCATGCCCATGGTGCAGCAGTGACACGCTTCACTACCCCTGGCTTCCCC 764
```

```
Db       2483 AGCTTTGGCCTGCACGCCCGCGGTGTGGAGCTGATGCGCTTCACCACGCCCGGCTTCCCT 2424

Qy         765 AACAGTCCCTACCCGGCGCATGCCCGCTGCCAGTGGGTCCTGCGGGGGGACGCCGACTCT 824
Db       2423 GACAGCCCCTACCCCGCTCATGCCCGCTGCCAGTGGGCCCTGCGGGGGGACGCCGACTCA 2364

Qy         825 GTGCTGAGCCTCACCTTCCGAAGCTTTGATGTCGCTCCCTGTGATGAGCATGGCAGTGAC 884
Db       2363 GTGCTGAGCCTCACCTTCCGCAGCTTTGACCTTGCGTCCTGCGACGAGCGCGGCAGCGAC 2304

Qy         885 CTGGTCACCGTGTATGATAGCCTGAGCCCCATGGAACCCCACGCTGTGGTGCGGCTGTGT 944
Db       2303 CTGGTGACGGTGTACAACACCCTGAGCCCCATGGAGCCCCACGCCCTGGTGCAGTTGTGT 2244

Qy         945 GGCACCTTCTCACCCTCCTACAACCTGACTTTCCTCTCCTCCCAGAACGTCTTCCTTGTC 1004
Db       2243 GGCACCTACCCTCCCTCCTACAACCTGACCTTCCACTCCTCCCAGAACGTCCTGCTCATC 2184

Qy        1005 ACGCTGATAACCAATACTGACCGGCGACATCCTGGCTTTGAGGCCACTTTCTTCCAGCTG 1064
Db       2183 ACACTGATAACCAACACTGAGCGGCGGCATCCCGGCTTTGAGGCCACCTTCTTCCAGCTG 2124

Qy        1065 CCCAAGATGAGCAGCTGTGGCGGCTTTTTGAGTGACACCCAAGGGACATTTAGCAGCCCC 1124
Db       2123 CCTAGGATGAGCAGCTGTGGAGGCCGCTTACGTAAAGCCCAGGGGACATTCAACAGCCCC 2064

Qy        1125 TACTATCCAGGCCACTACCCGCCCAACATCAACTGCACATGGAATATCAAGGTGCCCAAC 1184
Db       2063 TACTACCCAGGCCACTACCCACCCAACATTGACTGCACATGGAACATTGAGGTGCCCAAC 2004

Qy        1185 AACCGGAACGTGAAGGTGCGCTTCAAACTCTTCTATCTGGTGGACCCCAACGTACCAGTG 1244
Db       2003 AACCAGCATGTGAAGGTGAGCTTCAAATTCTTCTACCTGCTGGAGCCCGGCGTGCCTGCG 1944

Qy        1245 GGCTCCTGCACCAAGGACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGGTCC 1304
Db       1943 GGCACCTGCCCCAAGGACTACGTGGAGATCAATGGGGAGAAATACTGCGGAGAGAGGTCC 1884

Qy        1305 CAGTTTGTGGTGAGCAGCAACAGCAGCAAGATTACAGTCCACTTCCATTCTGATCACTCG 1364
Db       1883 CAGTTCGTCGTCACCAGCAACAGCAACAAGATCACAGTTCGCTTCCACTCAGATCAGTCC 1824

Qy        1365 TACACGGACACCGGGTTCCTAGCTGAGTACCTCTCCTACGACTCCAACGACCCGTGCCCA 1424
Db       1823 TACACCGACACCGGCTTCTTAGCTGAATACCTCTCCTACGACTCCAGTGACCCATGCCCG 1764

Qy        1425 GGGATGTTCATGTGCAAGACTGGACGGTGCATCCGAAAGGAACTGCGCTGCGACGGCTGG 1484
Db       1763 GGGCAGTTCACGTGCCGCACGGGGCGGTGTATCCGGAAGGAGCTGCGCTGTGATGGCTGG 1704

Qy        1485 GCAGACTGCCCGGATTATAGTGATGAGCGTTACTGCCGATGCAATGCCACCCACCAGTTC 1544
Db       1703 GCCGACTGCACCGACCACAGCGATGAGCTCAACTGCAGTTGCGACGCCGGCCACCAGTTC 1644

Qy        1545 ACGTGCAAAAACCAGTTCTGCAAGCCCCTCTTCTGGGTCTGTGACAGTGTCAACGACTGT 1604
Db       1643 ACGTGCAAGAACAAGTTCTGCAAGCCCCTCTTCTGGGTCTGCGACAGTGTGAACGACTGC 1584

Qy        1605 GGGGACGGAAGTGACGAGGAGGGCTGCAGCTGTCCTGCTGGGAGTTTCAAGTGTTCCAAT 1664
Db       1583 GGAGACAACAGCGACGAGCAGGGGTGCAGTTGTCCGGCCCAGACCTTCAGGTGTTCCAAT 1524

Qy        1665 GGGAAGTGTCTCCCTCAGAGCCAGAAGTGTAATGGGAAGGACAACTGTGGAGATGGGTCT 1724
Db       1523 GGGAAGTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTGGGGACGGGTCC 1464

Qy        1725 GACGAGGCTTCATGTGACAGCGTGAATGTCGTCTCTTGCACCAAATATACCTACCGCTGC 1784
Db       1463 GACGAGGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCTACCGCTGC 1404

Qy        1785 CAAAATGGCCTCTGTCTGAGCAAGGGCAACCCTGAGTGTGATGGGAAGACGGACTGTAGC 1844
Db       1403 CTCAATGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGGAGGACTGTAGC 1344
```

```
Qy       1845 GATGGCTCCGATGAGAAAAACTGTGACTGTGGGCTGCGATCCTTTACCAAACAGGCTCGC 1904
Db       1343 GACGGCTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGTCATTCACGAGACAGGCTCGT 1284

Qy       1905 GTGGTTGGTGGCACGAATGCGGACGAGGGCGAGTGGCCCTGGCAGGTGAGCCTCCACGCC 1964
Db       1283 GTTGTTGGGGGCACGGATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAAGCCTGCATGCT 1224

Qy       1965 CTGGGCCAGGGCCACTTGTGTGGGGCCTCGCTCATCTCTCCTGACTGGCTGGTCTCTGCA 2024
Db       1223 CTGGGCCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGCTGGTCTCTGCC 1164

Qy       2025 GCTCATTGCTTTCAGGATGACAAAAATTTCAAGTACTCAGACTACACGATGTGGACGGCC 2084
Db       1163 GCACACTGCTACATCGATGACAGAGGATTCAGGTACTCAGACCCCACGCAGTGGACGGCC 1104

Qy       2085 TTCCTGGGTCTGCTGGACCAGAGCAAGCGCAGTGCCTCTGGGGTGCAGGAGCTGAAGCTC 2144
Db       1103 TTCCTGGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCAGGAGCGCAGGCTC 1044

Qy       2145 AAACGTATCATCACCCACCCTTCCTTCAATGATTTCACCTTCGACTATGACATCGCCTTG 2204
Db       1043 AAGCGCATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATGACATCGCGCTG 984

Qy       2205 CTGGAGCTGGAGAAGTCGGTGGAGTACAGCACCGTCGTGCGCCCCATCTGCCTGCCTGAT 2264
Db        983 CTGGAGCTGGAGAAACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGCCGGAC 924

Qy       2265 GCTACCCATGTCTTCCCTGCTGGCAAGGCCATCTGGGTCACAGGCTGGGGGCACACAAAA 2324
Db        923 GCCTCCCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGGGCTGGGGACACACCCAG 864

Qy       2325 GAGGGAGGTACCGGAGCGCTGATCCTGCAGAAGGGTGAGATCCGTGTCATCAACCAGACC 2384
Db        863 TATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACC 804

Qy       2385 ACCTGTGAGGACCTCATGCCGCAGCAGATCACCCCACGAATGATGTGTGTGGGTTTCCTC 2444
Db        803 ACCTGCGAGAACCTCCTGCCGCAGCAGATCACGCCGCGCATGATGTGCGTGGGCTTCCTC 744

Qy       2445 AGTGGGGGTGTGGACTCCTGCCAGGGTGACTCTGGTGGCCCCTTGTCAAGCGCGGAGAAA 2504
Db        743 AGCGGCGGCGTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGTGGAGGCG 684

Qy       2505 GATGGGCGAATGTTCCAGGCTGGTGTGGTGAGCTGGGGTGAAGGCTGCGCTCAGAGGAAC 2564
Db        683 GATGGGCGGATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAAC 624

Qy       2565 AAGCCAGGCGTGTACACAAGGCTCCCTGTAGTTCGGGACTGGATCAAAGAGCACACTGGG 2624
Db        623 AAGCCAGGCGTGTACACAAGGCTCCCTCTGTTTCGGGACTGGATCAAAGAGAACACTGGG 564

Qy       2625 GTATAGCAGCATGGACAGACAGCCGACCACAAACACCCACAGGGATGCCCGACATGCACA 2684
Db        563 GTATAGGGGCCGGG---GCCACCCAAATGTGTACACCTGCGGGGCCACCCATCGTCCACC 507

Qy       2685 CCTGGATACAGGAGAGGAACACTGACGACATTTATGCTGTGGCCTCCCCCCCCCAACACA 2744
Db        506 CCAGTGTGCACG-CCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCC-CCAGA 449

Qy       2745 ACCCAGACTGTGAACTGCATCCTTAGGACTCAGAGT 2780
Db        448 ACATACACTGTGAACTCAATCTCCAGGGCTCCAAAT 413
```

Search completed: September 30, 2004, 14:24:58
Job time : 11831 secs

BLANK
PAGE